

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 54.433 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-3

Perfect score: 89

Sequence: 1 TYFPNKGTYQYTDQIE 16

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	370	12	AA13406
2	89	100.0	543	12	AA13405
3	89	100.0	554	16	AAW08987
4	89	100.0	554	20	AAW23230
5	89	100.0	554	24	ABP57264
6	89	100.0	554	24	ABP57267
7	89	100.0	756	21	AAV71231
8	89	100.0	781	16	AAW08986
9	89	100.0	781	20	AAV23227

10	89	100.0	781	24	ABP57263
11	89	100.0	781	24	ABP57266
12	45	50.6	767	22	ABG19946
13	44	49.4	904	22	ABG13993
14	42	47.2	1495	23	ABBS7380
15	42	47.2	1501	16	AA172858
16	42	47.2	1904	23	ABBS7100
17	42	47.2	1911	16	AA171726
18	42	47.2	1911	18	AAW27225
19	42	47.2	1911	20	AAW94027
20	42	47.2	1911	22	AAU01459
21	41	46.1	237	22	ABBS6093
22	41	46.1	283	22	AAU37883
23	41	46.1	393	23	ABBS6197
24	41	46.1	485	23	AA124847
25	41	46.1	2951	22	ABBS6291
26	40.5	45.5	349	23	ABBS5648
27	40	44.9	67	22	ABG26288
28	40	44.9	91	22	ABG16273
29	40	44.9	135	22	ABG10306
30	40	44.9	166	22	ABG10295
31	40	44.9	187	20	AA14929
32	40	44.9	250	22	ABG24534
33	40	44.9	283	19	AA186094
34	40	44.9	283	24	ABU02145
35	40	44.9	447	22	ABBS6406
36	40	44.9	450	21	AA11503
37	40	44.9	527	21	AA120291
38	40	44.9	577	21	AA120290
39	40	44.9	642	21	AA11505
40	40	44.9	645	21	AA11505
41	40	44.9	682	21	AA120289
42	40	44.9	692	21	AA11532
43	40	44.9	859	23	ABBS3471
44	40	44.9	869	23	ABBS6213
45	40	44.9	869	24	ABU66567

ALIGNMENTS

RESULT 1	AA13406	standard; Protein; 370 AA.
AA13406	AA13406	
AC	AA13406;	
XX		
DT	24-OCT-1991 (first entry)	
XX		
DE	Parvo virus B19 PANSE.	
XX		
KW	Primer; PCR; PAPST; globulin.	
XX		
OS	Synthetic.	
PN	DE4003826-A.	
XX		
PD	14-AUG-1991.	
XX		
XX		
PF	08-FEB-1990; 90DE-4003826.	
XX		
PR	08-FEB-1990; 90DE-4003826.	
XX		
PA	(MIKRO-) MIKROGEN MOLEKULARB.	
XX		
PI	Schutsek E, Motz M;	
XX		
DR	WPI; 1991-246423/34.	
PT	Immunologically active parvo virus B19 peptide(s) - comprising	
PT	capsid protein VP1 or VP2 fragments, useful for antibody	
PT	detection or vaccination	
XX		

Human parvovirus B
Human parvovirus B
Novel human diagno
Novel human diagno
Rat mucocardiatal ce
Rat receptor type-
Mouse ischaemic co
Human PTP-OB. Hom
Human protein tyro
Human protein tyro
Human protein tyro
Streptococcus pneu
Drosophila melanog
Infectious hematop
Drosophila melanog
Lactococcus lactis
Novel human diagno
Novel human diagno
Novel human diagno
Novel human diagno
Amino acid sequenc
Novel human diagno
S. pneumoniae deri
Drosophila melanog
SEN virus protein
Arabidopsis thailia
Arabidopsis thailia
Human ORF1320
Arabidopsis thailia
SEN virus protein
Human protein MDDT
HIV Env isolate TV
Human immunodefici

PS Claim 16; Page 10-11; 22pp; German.
 XX
 CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAQ13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 CC
 SQ Sequence 370 AA;

Query Match 100.0%; Score 89; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYFPNKGTOQYTDQIE 16
 |||||
 227 TYFPNKGTOQYTDQIE 242

RESULT 2
 AAR13405
 ID AAR13405 standard; Protein; 543 AA.
 XX
 AC AAR13405;
 XX
 DT 24-OCT-1991 (first entry)
 DE Parvo virus B19 VP2.
 XX
 DE Parvo virus B19 VP2.
 XX
 KM Primer; PCR; globulin; PANSE; PAPST.
 XX
 OS Synthetic.
 XX
 PN DE4003826-A.
 XX
 PD 14-AUG-1991.
 XX
 PF 08-FEB-1990; 90DE-4003826.
 XX
 PR 08-FEB-1990; 90DE-4003826.
 XX
 PA (MIKR-) MIKROGEN MOLEKULARB.
 XX
 XX Soutschek E, Motz M;
 XX
 WP1; 1991-246423/34.
 XX
 PT Immunologically active parvo virus B19 peptide(s) - comprising
 PT capsid protein VP1 or VP2 fragments, useful for antibody
 PT detection or vaccination
 PT
 XX
 PS Disclosure; Fig 2-6; 22pp; German.
 XX
 CC VP2 and its fragments PANSE (AAR13406) and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAQ13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 CC
 SQ Sequence 543 AA;

Query Match 100.0%; Score 89; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 |||||
 DB 400 TYFPNKGTOQYTDQIE 415

RESULT 3
 AAM08987
 ID AAM08987 standard; Protein; 554 AA.
 XX
 AC AAM08987;
 XX
 DT 27-FEB-1997 (first entry)
 DE Human parvovirus VP-2 protein.
 XX
 DE Human parvovirus VP-2 protein.
 XX
 KM Human; parvovirus genome; structural gene; VP-1, VP2; arthritis;
 KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KM erythblastemia; abortion; universal fetal hydrops; liver disease;
 KM haemorrhagic fever; rheumatism; detection; IgG antibody.
 XX
 OS Human parvovirus.
 XX
 PN JP07147986-A.
 XX
 PD 13-JUN-1995.
 XX
 PF 24-SEP-1992; 92JP-0281017.
 XX
 PR 24-SEP-1992; 92JP-0281017.
 XX
 PA (DENK-) DENKA SEIKEN KK.
 XX
 PA (ELED) DENKI KAGAKU KOGYO KK.
 XX
 DR WP1; 1995-242756/32.
 DR N-PSDB; AAT9535.
 XX
 XX Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.
 XX
 PS Claim 3; Page 7-9; 38pp; English.
 XX
 CC The sequences given in AAM08986 represent the parvovirus structural
 CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.
 CC
 SQ Sequence 554 AA;

Query Match 100.0%; Score 89; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 |||||
 DB 391 TYFPNKGTOQYTDQIE 406

RESULT 4
 AAY23230
 ID AAY23230 standard; Protein; 554 AA.
 XX
 AC AAY23230;
 XX
 DT 26-AUG-1999 (first entry)
 DE Erythrovirus V9 VP2 protein.
 XX
 DE Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KW erythrovirus screening; typing; immunoassay; VP2 protein.

XX Erythrovirus.
 OS FR2771751-A1.
 XX
 PD 04-JUN-1999.
 XX
 PF 03-DEC-1997; 97FR-0015197.
 XX
 PR 03-DEC-1997; 97FR-0015197.
 XX
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.
 XX
 PI Auguste V, Garbarg CA, Nguyen QT;
 DR WPI: 1999-349543/30.
 DR N-PSDB; AB259577.
 XX
 PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
 diagnosis of its infections
 Claim 19; Page 57-58; 80pp; French.

XX The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AB259577) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.

XX Sequence 554 AA;

Query Match 100.0%; Score 89; DB 20; Length 554;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTFPNKGTQGYTDQIE 16
 |||||
 DB 391 TTFPNKGTQGYTDQIE 406

RESULT 5
 ABP57264
 ID ABP57264 standard; Protein; 554 AA.
 XX
 AC ABP57264;
 XX
 DT 22-APR-2003 (first entry)
 DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.
 XX
 KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 PR 19-MAR-2002; 2002US-365956P.
 PR 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuantes S, Shyamala V;
 DR WPI: 2003-201510/19.
 DR N-PSDB; AB259574.
 PT

XX Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -
 XX
 PS Example 4; Fig 7B; 148pp; English.

XX The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to
 CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.

XX Sequence 554 AA;

Query Match 100.0%; Score 89; DB 24; Length 554;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTFPNKGTQGYTDQIE 16
 |||||
 DB 391 TTFPNKGTQGYTDQIE 406

RESULT 6
 ABP57267
 ID ABP57267 standard; Protein; 554 AA.
 XX
 AC ABP57267;
 XX
 DT 22-APR-2003 (first entry)
 DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
 XX
 KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 PR 19-MAR-2002; 2002US-365956P.
 PR 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuantes S, Shyamala V;
 DR WPI: 2003-201510/19.
 DR N-PSDB; AB259577.
 DR
 PT Detecting a human parvovirus B19 infection in a biological sample to

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic acid with a primer complementary to the 3'-terminal portion of the RNA target sequence -

XX Example 4; Fig 10B; 148bp; English.

XX The present invention describes a method for detecting a human parvovirus B19 infection in a biological sample. The method comprises reacting the isolated parvovirus B19 nucleic acid with a first oligonucleotide consisting of a first primer containing a complexing sequence sufficiently complementary to the 3'-terminal portion of the RNA target sequence to complex with. Also described: (1) amplifying a target parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer consisting of a promoter region recognised by a DNA-dependent RNA polymerase operably linked to a human parvovirus B19-specific complexing sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked to an acridinium ester label; and (6) a diagnostic test kit comprising an oligonucleotide primer of (4), and instructions for conducting the diagnostic test. The method is useful for detecting parvovirus infection in a biological sample, such as in blood products, to prevent transmission of the virus through blood and plasma derivatives or by close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267 represent sequences used in the exemplification of the present invention.

XX Sequence 554 AA;

SO Query Match 100.0%; Score 89; DB 24; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPNKGTCQYTDQIE 16
|||||
Db 391 TYPNKGTCQYTDQIE 406

RESULT 7
AAV71231
ID AAV71231 standard; Protein; 756 AA.

XX AAV71231;

DT 08-SEP-2000 (first entry)

XX Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.

XX Adeno-associated virus; AAV2; human parvovirus B19; chimeric; recombinant parvoviral vector; cellular tropism; cap protein; capsid; gene delivery; gene therapy; VP1; VP2; VP3.

XX Chimeric - Adeno associated virus serotype 2.

XX Chimeric - Human parvovirus B19.

XX WO200028004-A1.

XX PD 18-MAY-2000.

XX PF 10-NOV-1999; 99WO-US26505.

XX PR 10-NOV-1998; 98US-0107840.

XX PR 10-MAR-1999; 99US-0123651.

XX PA (UYN-) UNITV NORTH CAROLINA.

XX PI Rabinowitz JE, Samuleki RJ, Xiao W;

XX WPI: 2000-376523/32.

XX DR N-PSDB; AAD00833.

XX Recombinant parvoviral vectors with altered packaging, tropisms and

PT immunogenic properties, useful in gene therapy protocols -

XX Example 21; Page 142; 153bp; English.

XX The patent discloses modified parvovirus vectors with advantageous antigenic properties, packaging capabilities and cellular tropisms. These vectors can be used in standard recombinant DNA protocols e.g. gene therapy for delivering nucleic acids to cells.

XX The present sequence is a capsid protein encoded by an adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric vector. This vector encodes AAV2 VP1 and VP2 capsid proteins and human parvovirus B19 VP2 protein. The chimeric vector was constructed by replacing the VP3 major cap protein of AAV2 with B19's VP2. Recombinant parvovirus comprising the chimeric capsid is useful for gene delivery.

XX Sequence 756 AA;

SO Query Match 100.0%; Score 89; DB 21; Length 756;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYPNKGTCQYTDQIE 16
|||||
Db 593 TYPNKGTCQYTDQIE 608

RESULT 8
AAW08986
ID AAW08986 standard; Protein; 781 AA.

XX AAW08986;

DT 27-FEB-1997 (first entry)

XX Human parvovirus VP-1 protein.

XX Human; parvovirus genome; structural gene; VP-1; VP2; arthritis; non-structural protein; NS; diagnosis; vaccine; parvoviral disease; erythblastemia; abortion; universal fetal hydrops; liver disease; haemorrhagic fever; rheumatism; detection; IgG antibody.

XX Human parvovirus.

XX Key Location/Qualifiers

FT Misc-difference 19

FT FT /label= Gly, Val

FT FT Misc-difference 61

FT FT /label= Asn, Asp

FT FT Misc-difference 220

FT FT /label= His, Asn

PN JP07147986-A.

XX 13-JUN-1995.

XX PF 24-SEP-1992; 92JP-0281017.

XX PR 24-SEP-1992; 92JP-0281017.

XX PA (DENK-) DENKA SEIKEN KK.

XX PA (ELBD) DENKI KAKAKU KOGYO KK.

XX DR WPI: 1995-242756/32.

XX DR N-PSDB; AAT49535.

XX Human parvovirus gene coding for a polypeptide - useful for developing vaccines against parvoviral diseases such as erythroblastemia, haemorrhagic fever, etc.

XX Claim 2; Page 5-7; 38pp; English.

XX The sequences given in AAW08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

XX Sequence 781 AA;

Query Match 100.0%; Score 89; DB 16; Length 781;

Best Local Similarity 100.0%; Pred. No. 4e-06; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIE 16
 |||||
 Db 618 TYFPNKGTOQYTDQIE 633

RESULT 9

AAV23227 standard; Protein; 781 AA.

AAV23227;

26-AUG-1999 (first entry)

Erythrovirus V9 VP1 protein.

Erythrovirus V9; differential diagnosis; parvovirus; infection;

Erythrovirus screening; typing; immunoassay; VP1 protein.

Erythrovirus.

FR2771751-A1.

04-JUN-1999.

03-DEC-1997; 97FR-0015197.

03-DEC-1997; 97FR-0015197.

(ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.

Auguste V, Garbary CA, Nguyen QT;

WPI: 1999-349543/30.

N-PSDB; AAX81583.

Erythrovirus V9 and its nucleic acid sequences - can be used in the

diagnosis of its infections

Claim 19; Page 50-52; 80pp; French.

The present sequence represents an erythrovirus V9 protein.

CC probes and primers derived from erythrovirus V9 polynucleotide

sequences (AAX81580) can be used for differential diagnosis of

erythrovirus (parvovirus) infections by a combination of

amplification and hybridisation assay. The probes can also be

used to assess susceptibility to erythrovirus infection and

for erythrovirus screening and typing. The antibodies can be

used in immunoassays for diagnosis of erythrovirus V9 infections.

RESULT 10

ABP57263 standard; Protein; 781 AA.

ABP57263;

22-APR-2003 (first entry)

Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

Human parvovirus B19.

WO2003002753-A2.

09-JAN-2003.

28-JUN-2002; 2002WO-US20684.

28-JUN-2001; 2001US-302077P.

19-MAR-2002; 2002US-365956P.

29-MAR-2002; 2002US-369224P.

(CHIR) CHIRON CORP.

Pichuanes S, Shyamala V;

WPI: 2003-201510/19.

N-PSDB; AB259573.

Detecting a human parvovirus B19 infection in a biological sample to

prevent viral transmission, comprises reacting a parvovirus B19 nucleic

acid with a primer complementary to the 3'-terminal portion of the RNA

target sequence -

Example 4; Fig 6B; 148pp; English.

The present invention describes a method for detecting a human parvovirus

B19 infection in a biological sample. The method comprises reacting the

isolated parvovirus B19 nucleic acid with a first oligonucleotide

consisting of a first primer containing a complexing sequence

sufficiently complementary to the 3'-terminal portion of the RNA target

sequence to complex with. Also described: (1) amplifying a target

parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one

of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to

AB259629); (3) a polynucleotide comprising either of 2 4678 base pair

sequences (see AB259570 and AB259571); (4) an oligonucleotide primer

consisting of a promoter region recognised by a DNA-dependent RNA

polymerase operably linked to a human parvovirus B19-specific complexing

sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a

parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked

to an acridinium ester label; and (6) a diagnostic test kit comprising an

oligonucleotide primer of (4), and instructions for conducting the

diagnostic test. The method is useful for detecting parvovirus infection

in a biological sample, such as in blood products, to prevent

transmission of the virus through blood and plasma derivatives or by

close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267

represent sequences used in the exemplification of the present invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 89; DB 24; Length 781;

Best Local Similarity 100.0%; Pred. No. 4e-06; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIE 16
 |||||
 Db 618 TYFPNKGTOQYTDQIE 633

RESULT 11

ABP57266

ID ABP57266 standard; Protein; 781 AA.
 AC ABP57266;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B6-VP1 amino acid sequence SEQ ID NO:33.
 XX
 KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 XX
 PR 19-MAR-2002; 2002US-365956P.
 XX
 PR 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX
 DR WPI; 2003-201510/19.
 XX
 DR N-PSDB; ABZ59576.
 XX
 PT Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -
 XX
 PS Example 4; Fig 9B; 148pp; English.
 XX
 CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX
 SO Sequence 781 AA;
 Query Match 100.0%; Score 89; DB 24; Length 781;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTFPNKGTQOQYTDQIE 16
 |||||
 DB 618 TTFPNKGTQOQYTDQIE 633

RESULT 12
 ABG13946
 ID ABG13946 standard; Protein; 767 AA.
 AC
 XX

AC ABG13946;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19937.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 XX
 DR N-PSDB; AAS84133.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostic, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 50305; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 767 AA;
 Query Match 50.6%; Score 45; DB 22; Length 767;
 Best Local Similarity 53.8%; Pred. No. 89;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 3 FPNKGTQOQYTDQI 15
 ||:|:|:|:|:|
 DB 511 FPNKGTQYADQM 523

RESULT 13
 ABG13993
 ID ABG13993 standard; Protein; 904 AA.
 AC
 XX
 AC ABG13993;
 XX

DT 18-FEB-2002 (first entry)
 XX
 XX Novel human diagnostic protein #13984.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM Food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.
 DR N-PSDB; AAS78180.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostic, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID NO 44352; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 904 AA;

Query Match 49.4%; Score 44; DB 22; Length 904;
 Best Local Similarity 46.7%; Pred. No. 1.6e+02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYPNKGITQYTDQI 15
 ||||| : : :
 DB 168 TYAPNKGASRPIKQV 162

RESULT 14
 ABB57380
 ID ABB57380 standard; Protein; 1495 AA.
 XX
 AC ABB57380;
 XX
 DT 08-MAR-2002 (first entry)
 XX

DB Rat myocardial cell proliferation associated polypeptide SEQ ID NO 12.
 XX
 XX Rat; heart; cardiac; myocardial necrosis; cardiac hypertrophy;
 KM cardiac insufficiency.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200183705-A1.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-JP03700.
 XX
 PR 27-APR-2000; 2000JP-0126741.
 XX

PA (KYOM) KYOMA HAKKO KOGYO KK.
 XX
 PI Yamada Y, Sekine S, Kikuchi Y, Sakurada K;
 XX
 DR WPI: 2002-075160/10.
 DR N-PSDB; AB199920.

PT Genes having differential expression in fetal and adult heart tissue
 PT useful for screening potential drugs for promoting repair of damage
 PT caused by myocardial necrosis -
 PS Claim 53; Page 110-116; 171bp; Japanese.

CC The invention relates to gene sequences (AB199915-AB199934) having
 CC modified expression in fetal heart tissue as compared to adult heart
 CC tissue and the encoded proteins (ABB57375-ABB57392). The genes have
 CC cardiac activity and may be useful in the promotion of the repair of
 CC damage to heart tissue caused by myocardial necrosis. The gene sequences
 CC are useful for screening potential compounds for the ability to influence
 CC disease associated with myocardial necrosis. Drugs identified by the
 CC screening methods may be used to treat and prevent disease with which
 CC myocardial necrosis is associated, such as cardiac hypertrophy and
 CC cardiac insufficiency. Diagnosis of diseases such as those above is also
 CC disclosed.

CC Sequence 1495 AA;

Query Match 47.2%; Score 42; DB 23; Length 1495;
 Best Local Similarity 60.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGITQY 11
 : : : : :
 DB 1054 YFPNKGITQY 1063

RESULT 15
 AAR72858
 ID AAR72858 standard; Protein; 1501 AA.
 XX
 AC AAR72858;
 XX
 DT 25-MAR-2003 (updated)
 DT 21-NOV-1995 (first entry)
 XX
 DE Rat receptor type-tyrosine phosphatase sigma.
 XX
 KM Receptor type tyrosine phosphatase sigma; cell differentiation;
 KM metabolism; cell cycle; behaviour; motility; contact inhibition;
 KM virus; inflammation; cellular transformation; cancer;
 KM neuroblastomas; antibody; detection; quantification.
 XX
 OS Rattus rattus.
 XX
 PN WO9509656-A1.
 XX
 PD 13-APR-1995.
 XX

PF 30-SEP-1994; 94WO-US11163.
 XX
 PR 01-OCT-1993; 93US-0130570.
 XX
 PA (UTNY) UNIV NEW YORK STATE.
 XX
 PI Schlessinger J, Van H;
 XX
 DR WPI; 1995-155068/20.
 DR N-PSDB; AAO86902.
 XX
 PT Novel, isolated receptor-tyrosine protein tyrosine phosphatase-sigma
 PT - and encoding DNA, useful e.g. for detecting neuro-blastomas
 XX
 PS Claim 2; Figure 2; 105pp; English.
 XX
 CC Ligands binding to the receptor-tyrosine protein tyrosine phosphatase
 CC sigma (RTP sigma) protein may be used as drugs to modulate cellular
 CC processes, such as differentiation, metabolism and cell cycle
 CC control, and cellular behaviour such as motility and contact
 CC inhibitions. In addition they may affect abnormal or potentially
 CC deleterious processes such as virus-receptor interactions,
 CC inflammation and cellular transformation to a cancerous state. They
 CC may also be used to treat RTP sigma related neuronal disorders such
 CC as neuroblastomas. The DNA encoding the RTP sigma is useful for
 CC the diagnosis of diseases resulting from its aberrant expression.
 CC Antibodies directed against RTP sigma may be used in detection and
 CC quantitative analysis
 CC (updated on 25-MAR-2003 to correct PN field.)
 CC
 SQ Sequence 1501 AA;

Query Match 47.2%; Score 42; DB 16; Length 1501;
 Best Local Similarity 60.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 2 YFPNKGTOOY 11
 Db 1057 YMPNKGTEY 1066

RESULT 16
 ABB57100
 ID ABB57100 standard; Protein; 1904 AA.
 XX
 AC ABB57100;
 XX
 OY 07-MAR-2002 (first entry)

Mouse ischaemic condition related protein sequence SEQ ID NO:224.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.
 XX

OS Mus musculus.

XX WO200186188-A2.

XX PD 22-NOV-2001.

XX PF 18-MAY-2001; 2001WO-JP04192.

XX PR 18-MAY-2000; 2000JP-0145977.

XX PA (UTNY-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

XX DR N-PSDB; AB199344.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or

PT by determining the expression profile of a gene group comprising these
 PT genes -
 XX
 PS Claim 2; Page 636-644; 2690pp; English.

XX The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (AB199202 to AB199912, encoding
 CC the protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.

SQ Sequence 1904 AA;
 Query Match 47.2%; Score 42; DB 23; Length 1904;
 Best Local Similarity 60.0%; Pred. No. 8.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 2 YFPNKGTOOY 11
 Db 1463 YMPNKGTEY 1472

RESULT 17
 AAR71726
 ID AAR71726 standard; Protein; 1911 AA.
 XX
 AC AAR71726;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-OCT-1995 (first entry)
 XX

DE Human PTP-OB.

XX PTP-OB; protein tyrosine phosphatase; osteoblast; differentiation;
 KM osteoclast; osteoporosis; bone; cancer; osteosarcoma.
 XX

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= Sig_peptide

FT Modified-site 250

FT /label= N-glycosylation_site

FT Modified-site 721

FT /label= N-glycosylation_site

FT Modified-site 919

FT /label= N-glycosylation_site

FT Domain 1253..1277

FT /label= Extracellular_domain

XX WO9507935-A1.

XX PN 23-MAR-1995.

XX PD 09-SEP-1994; 94WO-US10166.

XX PF 14-SEP-1993; 93US-0122032.

XX PR (MERI) MERCK & CO INC.

XX PA Rodan GA, Rutledge SJ, Schmidt A;

XX PI WPI; 1995-131318/17.

XX DR N-PSDB; AAO86478.

XX Protein tyrosine phosphatase protein PTP-OB specifically expressed
 PT in bone cells - modulators of which are used to treat, e.g.
 PR osteoporosis, and prevent and treat bone loss and cancer.
 PS Claim 1, Page 44-45; 63pp; English.

XX PCR amplification of cDNA derived from human osteosarcoma
 CC Saos-2/B10 using primers based on conserved regions of protein
 CC tyrosine phosphatases and subsequent screening of a human
 CC brain cDNA library yielded a cDNA clone (sequence given in
 CC AA086473) that encoded a novel human protein, PTP-OB (AA071726).
 CC Recombinant PTP-OB was expressed in E. coli, yeast, insect
 CC and mammalian cells.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1911 AA;
 SQ

Query Match 47.2%; Score 42; DB 16; Length 1911;
 Best Local Similarity 60.0%; Pred. No. 8.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 Db 1467 YWPNRGTETY 1476

RESULT 18
 AAW27225
 ID AAW27225 standard; Protein; 1911 AA.
 XX
 AC AAW27225;
 XX
 DT 25-MAR-2003 (updated)
 DT 19-DEC-1997 (first entry)

XX Human protein tyrosine phosphatase PTP-OB.
 DE
 XX Protein tyrosine phosphatase 'PTP-OB; PTPepsilon; osteoblast;
 KW recombinant protein; growth; differentiation; brain; human.
 XX
 OS Homo sapiens.
 XX
 PN US5658756-A.
 XX
 PD 19-AUG-1997.
 XX
 PS 01-DEC-1994; 94US-0348006.
 XX
 PR 01-DEC-1994; 94US-0348006.
 PR 14-SEP-1993; 93US-0122032.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Rodan GA, Rutledge SJ, Schmidt A;
 XX
 DR WPI; 1997-424232/39.
 DR N-PSDB; AAT85389.
 XX
 PT DNA encoding protein tyrosine phosphatase PTP-OB - isolated from
 PT human osteoblasts and useful for production of recombinant PTP-OB

XX Claim 1; Column 23-34; 34pp; English.

XX The present sequence represents human protein tyrosine phosphatase
 CC (PTP-OB) protein. The DNA encoding this protein is useful for the
 CC production of the recombinant protein, which is a protein tyrosine
 CC phosphatase which may be involved in the growth and differentiation
 CC of osteoblasts and brain cells and is useful for identifying compounds
 CC that modulate PTP-OB activity and as a therapeutic agent for treating
 CC PTP-OB-related diseases.
 CC (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 1911 AA;
 Query Match 47.2%; Score 42; DB 16; Length 1911;
 Best Local Similarity 60.0%; Pred. No. 8.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 Db 1467 YWPNRGTETY 1476

RESULT 19
 AAW94027
 ID AAW94027 standard; Protein; 1911 AA.
 XX
 AC AAW94027;
 XX
 DT 01-APR-1999 (first entry)
 DT
 DE Human protein tyrosine phosphatase (PTP-OB).
 XX
 KW Protein tyrosine phosphatase; PTP; PTP-OB; bone; brain; cancer;
 KW osteoporosis.
 XX
 OS Homo sapiens.
 XX
 PN US5866397-A.
 XX
 PD 02-FEB-1999.
 XX
 PF 14-FEB-1997; 97US-0800825.
 XX
 PR 01-DEC-1994; 94US-0348006.
 PR 14-SEP-1993; 93US-0122032.
 PR 14-FEB-1997; 97US-0800825.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Rodan GA, Rutledge SJ, Schmidt A;
 XX
 DR WPI; 1999-141930/12.
 DR N-PSDB; AAX06095.
 XX
 PT Protein tyrosine phosphatase denoted PTP-OB - useful for drug
 PT screening

XX Claim 1; Columns 23-32; 34pp; English.

XX This represents a human protein tyrosine phosphatase (PTP) denoted as
 CC PTP-OB, produced by bone and brain cells. A recombinant host cell
 CC transfected or transformed with a nucleic acid vector comprising the
 CC nucleic acid can be used for the production of the PTP-OB polypeptide.
 CC The protein can be used to screen for modulators of PTP-OB activity,
 CC which might be useful for treating e.g. osteoporosis and cancer.

XX Sequence 1911 AA;
 Query Match 47.2%; Score 42; DB 20; Length 1911;
 Best Local Similarity 60.0%; Pred. No. 8.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 Db 1467 YWPNRGTETY 1476

RESULT 20
 AAU01459
 ID AAU01459 standard; Protein; 1911 AA.
 XX
 AC AAU01459;
 XX
 DT 18-JUL-2001 (first entry)

XX Human protein tyrosine phosphatase, PTP-OB, amino acid sequence.
 DE
 XX
 KM Protein tyrosine phosphatase; PTP; human; osteoporosis; bone formation;
 KM PTP-OB; cell proliferation; cell death; neoplastic transformation;
 KM metastasis; tumour; cancer; lung cancer; osteosarcoma.
 OS
 XX Homo sapiens.
 XX
 PN US6214564-B1.
 PD
 XX 10-APR-2001.
 PD
 XX 22-SEP-1998; 98US-0158657.
 PF
 XX 01-DEC-1994; 94US-0348006.
 PR 14-FEB-1997; 97US-0800825.
 PR 14-SEP-1993; 93US-0122032.
 XX
 XX (MERI) MERCK & CO INC.

Rodan GA, Rutledge SJ, Schmidt A;

XX WPI; 2001-280994/29.
 DR N-PSDB; AAS02144.
 XX

PT Identifying a compound which modulates protein tyrosine phosphatase
 PT activity for treating osteoporosis, comprises contacting a polypeptide
 PT with a compound and measuring the ability of the compound to modulate
 PT polypeptide activity -
 PT
 XX
 PS Claim 1; Column 23-33; 35pp; English.

XX The sequence represents the amino acid sequence of human protein tyrosine
 CC phosphatase, PTP-OB. The PTP-OB nucleic acid and protein sequences
 CC were used in a method to identify a compound able to modulate PTP-OB
 CC activity, involving contacting a recombinant polypeptide comprising the
 CC entire cytoplasmic domain of PTP-OB with the compound, and measuring the
 CC ability of the compound to modulate the activity of the polypeptide. The
 CC compounds that modulate PTP-OB activity are useful in treating disease
 CC states involving PTP-OB activity such as osteoporosis, for preventing and
 CC treating bone loss, and stimulation of bone formation. The compounds are
 CC also useful for treating diseases in which activation or inactivation of
 CC PTP-OB protein results in either cellular proliferation, cell death,
 CC nonproliferation, induction of cellular neoplastic transformations or
 CC metastatic tumour growth, such as cancer, preferably lung cancer or
 CC osteosarcoma.

Sequence 1911 AA;

Query Match 47.2%; Score 42; DB 22; Length 1911;
 Best Local Similarity 60.0%; Pred. No. 8.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 |:|:|:|:|:
 Db 1467 YWPNRGHETV 1476

RESULT 21
 ABB60093
 ID ABB60093 standard; Protein; 237 AA.
 XX
 AC ABB60093;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 7071.

KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS Drosophila melanogaster.

XX WO200171042-A2.
 PN
 XX

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

XX WPI; 2001-656660/75.
 DR N-PSDB; ABL04196.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX

PS Disclosure; SEQ ID NO 7071; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
 CC sequences (ABBS7737-ABBS72072).
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 237 AA;

Query Match 46.1%; Score 41; DB 22; Length 237;
 Best Local Similarity 70.0%; Pred. No. 1e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 NKGTOOYTDQ 14
 |:|:|:|:|:
 Db 30 NNGTQKYTNQ 39

RESULT 22

ID AAU37883
 XX AAU37883 standard; Protein; 283 AA.

AC AAU37883;

DT 14-FEB-2002 (first entry)

XX Streptococcus pneumoniae cellular proliferation protein #312.

KM Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.

OS Streptococcus pneumoniae.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAB55742.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13476; 511pp; English.
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 283 AA;
 Query Match 46.1%; Score 41; DB 22; Length 283;
 Best Local Similarity 46.7%; Pred. No. 1.3e+02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 TYPNKGTOQYTDQI 15
 | | | | | : | | | :
 DB 205 TSPNKNSGQYTDQI 219
 | | | | | : | | | :
 RESULT 23
 AAB61197
 AAB61197 standard; Protein; 393 AA.
 AC ABB61197;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 10383.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 OS WO200171042-A2.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;

XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL05300.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 10383; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL016176-ABL030511), expressed DNA
 CC sequences (ABL016176-ABL016175) and the encoded proteins
 CC (AAB57737-ABR72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 393 AA;
 Query Match 46.1%; Score 41; DB 22; Length 393;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 YFPNKGTOQYTDQI 15
 | | | | | : | | | :
 DB 169 YFPLEGRQSYTPRM 182
 | | | | | : | | | :
 RESULT 24
 AAE24847
 ID AAE24847 standard; Protein; 485 AA.
 XX
 AC AAE24847;
 XX
 DT 30-OCT-2002 (first entry)
 XX
 DE Infectious hematopoietic necrosis virus-SCS G protein.
 XX
 DE G protein; M1 protein; infectious haematopoietic necrosis virus; IHNV;
 KW structural protein; vaccine; haematopoietic necrosis symptom; virulence;
 KW infectious hematopoietic necrosis virus; IHNV.
 XX
 OS Infectious hematopoietic necrosis virus.
 XX
 OS Key Location/Qualifiers
 FT Misc-difference 209 /note="Encoded by NGS"
 FT Misc-difference 228 /note="Encoded by ATN"
 XX
 PN WO200236618-A1.
 XX
 PD 10-MAY-2002.
 XX
 PF 05-NOV-2001; 2001WO-KR01871.
 XX
 PR 04-NOV-2000; 2000KR-0065384.
 XX
 PA (RMR-) RNA INC.
 XX
 PI Park J, Moon C, Yoon W, Park K, Kim H, Oh T, Kim H, Chung D;
 PI Cho W, Park J;
 XX
 DR WPI; 2002-519231/55.
 XX
 DR N-PSDB; AAD40506.
 XX
 PT New G or M1 structural proteins of a Korean isolate infectious
 PT hematopoietic necrosis virus (IHNV-SCS), for use in vaccines against

PT Infectious hematopoietic necrosis virus, particularly in fish, e.g.
 PT Salmonid fish -
 XX
 PS Claim 1; Page 54-56; 62pp; English.
 XX
 CC The present invention relates to novel polynucleotides encoding G or M1
 CC structural proteins of infectious haematopoietic necrosis virus (IHNV),
 CC which are isolated from Korean type IHNV-SCS. The G or M1 structural
 CC proteins of IHNV-SCS are useful as prophylactic vaccines against IHNV.
 CC The genes encoding the proteins are useful as DNA vaccines. The vaccines
 CC are particularly useful against haematopoietic necrosis symptoms in fish,
 CC e.g. Salmonid fish. The present sequence is IHNV-SCS G protein.
 XX
 SQ Sequence 485 AA;
 Query Match 46.1%; Score 41; DB 23; Length 485;
 Best Local Similarity 46.7%; Pred. No. 2,4e+02;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

2 YFPNKGTOQYTDQIE 16
 ||| ||| ||| |||
 144 YFPASRCQWYTDNVQ 158

RESULT 25
 ID ABB60291 standard; Protein; 2951 AA.
 XX ABB60291;
 AC ABB60291;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7665.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL04394.
 DR
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 7665; 21pp + Sequence listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB170511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 2951 AA;
 Query Match 46.1%; Score 41; DB 22; Length 2951;
 Best Local Similarity 63.6%; Pred. No. 2e+03;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PNNKGTQYTDQ 14
 ||| ||| ||| |||
 Db 506 PNNKGTQYTDQ 516

RESULT 26
 ID ABB53648 standard; Protein; 349 AA.
 XX ABB53648
 AC ABB53648;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein ddl.
 XX
 KM Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX WPI; 2002-043418/06.
 DR
 PT New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species -
 XX
 PS Claim 6; SEQ ID NO 350; 2504pp; French.
 XX
 CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 349 AA;

Query Match 45.5%; Score 40.5; DB 23; Length 349;
 Best Local Similarity 47.4%; Pred. No. 2e+02;
 Matches 9; Conservative 3; Mismatches 2; Indels 5; Gaps 1;

QY 1 TYFPNKG-----TOOYTDQ 14
 ||| ||| ||| |||
 Db 40 TYFITKGGEFTKQSTDK 58

RESULT 27
 ID ABB26288 standard; Protein; 67 AA.
 AC ABB26288;
 XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Dmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS74493.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 40665; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 SQ Sequence 135 AA;
 Query Match 44.9%; Score 40; DB 22; Length 135;
 Best Local Similarity 50.0%; Pred. No. 80;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 YFPNKGTOQYTDQI 15
 |||::|||
 Db 5 YFPHQPAQKXFQOI 18
 RESULT 30
 ABG10295
 ID ABG10295 standard; Protein; 166 AA.
 XX
 XX ABG10295;
 XX 13-FEB-2002 (first entry)
 XX Novel human diagnostic protein #10286.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
 OS WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Dmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS74482.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 40654; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 SQ Sequence 166 AA;
 Query Match 44.9%; Score 40; DB 22; Length 166;
 Best Local Similarity 50.0%; Pred. No. 16+02;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 YFPNKGTOQYTDQI 15
 |||::|||
 Db 5 YFPHQPAQKXFQOI 18
 RESULT 31
 AAY14929
 ID AAY14929 standard; Protein; 187 AA.
 XX
 XX AAY14929;
 XX 25-OCT-1999 (first entry)
 XX Amino acid sequence of M. vaccae antigen GV-41B.
 DE Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;

KM dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 KW squamous cell carcinoma; melanoma.
 OS Mycobacterium vaccae.
 XX WO9932634-A2.
 XX PD 01-JUL-1999.
 XX PF 23-DEC-1998; 98WO-NZ00189.
 XX PR 04-DEC-1998; 98US-0205426.
 XX PR 23-DEC-1997; 97US-0996624.
 XX PR 23-DEC-1997; 97US-0997080.
 XX PR 23-DEC-1997; 97US-0997362.
 XX PR 11-JUN-1998; 98US-0095855.
 XX PR 17-SEP-1998; 98US-0156181.
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX Prestidge RL, Skimer MA, Tan P, Visser ES, Watson J;
 DR WPI; 1999-430163/36.
 DR N-PSDB; AA211394.
 XX PT Enhancing immune response to an antigen
 XX Claim 1; Page 240; 243pp; English.
 XX PS
 CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.
 CC XX
 SQ Sequence 187 AA;
 ery Match 44.9%; Score 40; DB 20; Length 187;
 at Local Similarity 58.3%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 5 NKGITQYTDQIR 16
 :|:|:|:|:|:|:
 Db 161 DKSTHQYTNQID 172
 RESULT 32
 ABG24534
 ID ABG24534 standard; Protein; 250 AA.
 XX AC ABG24534;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #24525.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX PN WO200175067-A2.

XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX PF 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSB-) HYSB INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AA88721.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX PS Claim 20; SEQ ID No 54893; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX
 SQ Sequence 250 AA;
 Query Match 44.9%; Score 40; DB 22; Length 250;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 2 YFPNKGITQYTDQI 15
 :|:|:|:|:|:|:
 Db 5 YFPHQPAQKYFQOI 18
 RESULT 33
 AAY86094
 ID AAY86094 standard; Protein; 283 AA.
 XX AC AAY86094;
 XX DT 10-APR-2000 (first entry)
 XX DE S. pneumoniae derived protein #303.
 XX KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
 XX KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
 OS Streptococcus pneumoniae.
 XX PN WO9806734-A1.
 XX PD 19-FEB-1998.

XX 15-AUG-1997; 97MO-US14436.
 XX
 PR 16-AUG-1996; 96US-0024022.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 XX Scodola RK;
 XX
 DR MPI; 1998-159452/14.
 DR N-PSDB; AA296413.
 XX
 PT Streptococcus pneumoniae proteins and related DNA - useful for
 PT screening compounds for antibacterial activity
 XX
 PS Claim 5; Page 570; 640pp; English.

CC This invention describes novel isolated Streptococcus pneumoniae
 CC polynucleotides (see AA296173-296494) and their encoded proteins (see
 CC AA85792-Y86182). The DNA, vectors and host cells described in the
 CC method of the invention are useful for the recombinant expression of the
 CC polypeptides. The polypeptides are useful for treatment or prevention of
 CC disease, or diagnosis of disease related to expression or activity of
 CC such a polypeptide. They can also be used to screen for compounds which
 CC interact with and inhibit or activate such a polypeptide. The
 CC polypeptides (or DNA encoding them, via gene therapy) are also useful
 CC for inducing an immunological response in a mammal. The antagonists are
 CC useful to inhibit such bacterial polypeptides. The polypeptides are
 CC particularly useful to identify antimicrobial compounds and antibiotics.
 CC They are also useful to determine their role in pathogenesis of
 CC infection, dysfunction and disease.

SO Sequence 283 AA;

QY Query Match 44.9%; Score 40; DB 19; Length 283;
 Best Local Similarity 46.7%; Pred. NO. 1.9e+02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 DB 205 TSVPNKDSQTYTETV 219

RESULT 34
 ABU02145
 ID ABU02145 standard; Protein; 283 AA.

ABU02145;

DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #1722.

KW Bacterial meningitis; pneumonia; sepsis; otitis media;
 KM ear infection; antiinflammatory; antibacterial; immunostimulant;
 XX auditory; respiratory; gene therapy; vaccine.

OS Streptococcus pneumoniae type 4 strain.

PN WO200277021-A2.

PD 03-OCT-2002.

PF 27-MAR-2002; 2002WO-IB02163.

PR 27-MAR-2001; 2001GB-0007658.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Masignani V, Tettelin H, Fraser C;

DR MPI; 2003-040579/03.
 DR N-PSDB; ABX07434.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 PT or ear infection
 XX
 PS Claim 1; SEQ ID NO 3444; 56pp; English.

CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC AB55454. Also included are an antibody which binds one of the
 CC proteins, treating a patient by administering the protein, DNA or
 CC antibody (in a composition), a kit comprising first and second primers,
 CC which are the nucleic acid cited above or fragments between nucleotides
 CC 8-100 of a sequence not defined in the specification, for amplifying a
 CC target sequence contained within a Streptococcus nucleic acid sequence,
 CC where the first primer is substantially complementary to the target
 CC sequence and the second primer is substantially complementary to the
 CC complement of the target sequence, and where the parts of the primers
 CC having substantial complementarity define the termini of the target
 CC sequence to be amplified, assay comprising contacting a test compound
 CC with the protein, and determining whether the test compound binds to the
 CC protein and a Streptococcus pneumoniae bacterium, where one or more
 CC genes encoding the proteins has been rendered inactive. The protein,
 CC nucleic acid molecules, antibody and compositions are useful as
 CC medicaments for treating or preventing a disease or infection due to
 CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
 CC sepsis, otitis media or ear infection. They are also useful in developing
 CC vaccines, diagnostics and antibiotics. The methods are useful for
 CC identifying immunodominant proteins. The present sequence is one of
 CC the 2469 proteins expressed by the identified coding regions from the
 CC genomic sequence.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 283 AA;

QY Query Match 44.9%; Score 40; DB 24; Length 283;
 Best Local Similarity 46.7%; Pred. NO. 1.9e+02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

DB 205 TSVPNKDSQTYTETV 219

RESULT 35
 ABB65406
 ID ABB65406 standard; Protein; 447 AA.

ABB65406;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 23010.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI: 2001-656860/75.
DR N-PSDB; ABL09509.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 23010; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU01840-ABU16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 447 AA;
XX
Query Match 44.9%; Score 40; DB 22; Length 447;
Best Local Similarity 70.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 4 PNRGTQGYTD 13
DB 52 PNRGTQGYTD 61
XX
RESULT 36
AAB11503
ID AAB11503 standard; Protein; 450 AA.
XX
AC AAB11503;
XX
DT 12-OCT-2000 (first entry)
XX
SEN virus protein fragment SEQ ID NO: 10.
XX
KM SEN virus; SENV; gastrointestinal tract disorder; inflammatory disease;
KM proliferative disorder; hepatopathy; hepatitis; viral infection;
KM vaccination; gene therapy.
XX
OS Hepatitis virus.
XX
PN WO200028039-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-EP08566.
XX
PR 10-NOV-1998; 98IT-MI02437.
PR 30-APR-1999; 99IT-MI00923.
PR 14-MAY-1999; 99EP-0830298.
PR 16-JUL-1999; 99EP-0113932.
XX
PA (DIAS-) DIASORIN SRL.
XX
PI Print D, Fioridalisi G, Mantero GU, Mattioli S, Sottini A;
PI Bonelli F, Vaglini U, Olivero P, Dal Corso A, Bonelli M;
XX
DR MPI: 2000-376551/32.

XX Nucleic acids representing the genome of the SEN virus (SENV) and
PT encoded proteins; useful for treatment of hepatopathies, inflammatory
PT diseases and proliferative disorders such as cancer -
XX
PS Example 5; Page 299-300; 392pp; English.
XX
CC The present invention is concerned with the sequence of the genome of the
CC SEN virus (SENV), and the proteins encoded by it. SENV is thought to be
CC the cause of hepatopathies which are not linked to the presence of the
CC hepatitis A, B and E viruses in man. The genome and proteins of this
CC virus can be used in gene therapy and vaccination against the virus,
CC which also causes disorders of the gastrointestinal tract, including
CC Crohn's disease and lupus erythematosus, inflammatory diseases, and
CC proliferative disorders such as cancer.
XX
SQ Sequence 450 AA;
XX
Query Match 44.9%; Score 40; DB 21; Length 450;
Best Local Similarity 63.6%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
QY 3 FNRGTQGYTD 13
DB 152 FNRGTQGYTD 162
XX
RESULT 37
AAG20291
ID AAG20291 standard; Protein; 527 AA.
XX
AC AAG20291;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22419.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131445.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.9%; Score 40; DB 21; Length 527;
Best Local Similarity 52.6%; Pred. No. 3.9e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 3 PPNKGTQ-----QYTDQI 15
DB 89 PFGGSGTQFIHGADQYIDQI 107

ID AAG20290 standard; Protein; 577 AA.
XX AAG20290;
AC AAG20290;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22418.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144088.
PR 16-JUL-1999; 99US-0144088.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

```

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

```

```

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 44.9%; Score 40; DB 21; Length 577;
Best Local Similarity 52.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

```

```

QY 3 FPNKGTO-----QYTDQI 15
DB 139 FPGGTGFTHGADQYRDDI 157

```

```

RESULT 39
ID AAB11505 standard; Protein; 642 AA.
AAB11505

```

```

AC AAB11505;
XX
DT 12-OCT-2000 (first entry)
XX

```

```

DE SEN virus protein fragment SEQ ID NO: 25.
XX

```

```

KM SEN virus; SENV; gastrointestinal tract disorder; inflammatory disease;
proliferative disorder; hepatopathy; hepatitis; viral infection;
KM vaccination; gene therapy.
XX

```

```

OS Hepatitis virus.
XX

```

```

PN WO200028039-A2.
XX

```

```

PD 18-MAY-2000.
XX

```

```

PF 09-NOV-1999; 99WO-EP08566.
XX

```

```

PR 10-NOV-1998; 98IT-MI02437.
XX

```

```

PR 30-APR-1999; 99IT-MI00923.
XX

```

```

PR 14-MAY-1999; 99EP-0830298.
XX

```

```

PR 16-JUL-1999; 99EP-0113932.
XX

```

```

PA (DIAS-) DIASORIN SRL.
XX

```

```

PI Priml D, Fiordalisi G, Mantero GL, Mattioli S, Sottini A;
XX

```

```

PI Bonelli F, Vaglini L, Olivero P, Dal Corso A, Bonelli M;
XX

```

```

PT MPI; 2000-376551/32.
XX

```

```

PS Claim 1, Page 305-307; 392pp; English.
XX

```

```

XX The present invention is concerned with the sequence of the genome of the
XX SEN virus (SENV), and the proteins encoded by it. SENV is thought to be
XX the cause of hepatopathies which are not linked to the presence of the
XX hepatitis A, B and E viruses in man. The genome and proteins of this
XX virus can be used in gene therapy and vaccination against the virus,
XX which also causes disorders of the gastrointestinal tract, including
XX Crohn's disease and lupus erythematosus, inflammatory diseases, and
XX proliferative disorders such as cancer.
XX

```

```

SQ Sequence 642 AA;

```

```

Query Match 44.9%; Score 40; DB 21; Length 642;
Best Local Similarity 63.6%; Pred. No. 5e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 FPNKGTOQYTD 13
DB 180 FPMKGEAYTD 190

```


RESULT 40
 AAB41556
 ID AAB41556 standard; Protein; 645 AA.
 XX
 AC AAB41556;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1320 polypeptide sequence SEQ ID NO:2640.
 XX
 KM Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KM vulnery; antiparietal; antiparkinsonian; neurotropic; neuroprotective;
 KM anticonvulsant; osteopathic; antiairrhritic; immunosuppressant; cardiant;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antinflammatory;
 KM antiviral; antibacterial; antifungal; antineumatic; antihyroid;
 KM antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; nocturnal haemoglobinuria; burn; wound;
 KM thrombosis; contraceptive.
 KM
 KM Homo sapiens.
 OS
 XX
 PN WO20056473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 98US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkete RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC75765.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 Claim 11; Page 1885-1886; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparietal; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiairrhritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antineumatic; antihyroid;
 CC antihyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease, to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX

SQL Sequence 645 AA;
 Query Match 44.9%; Score 40; DB 21; Length 645;
 Best Local Similarity 46.2%; Pred. No. 5e+02; 3; Indels 0; Gaps 0;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 4 PNKGTOQYTDQIE 16
 | :|||:|
 DB 27 PTAASEQYTDRLR 39

Search completed: August 20, 2003, 09:29:35
 Job time : 56.433 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 17.6495 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-3

Perfect score: 89

Sequence: 1 TYFPNKGTYQYTDQIE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

1 number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	395	3	US-08-856-841-13
2	89	100.0	398	3	US-08-856-841-21
3	89	100.0	415	3	US-08-856-841-20
4	89	100.0	543	3	US-08-856-841-22
5	89	100.0	756	4	US-09-438-268-4
6	44	49.4	119	4	US-09-345-236B-135
7	44	49.4	235	4	US-09-345-236B-108
8	42	47.2	1501	2	US-08-447-464-3
9	42	47.2	1501	1	US-08-348-006B-5
10	42	47.2	1911	1	US-08-800-825A-5
11	42	47.2	1911	3	US-09-158-657-5
12	42	47.2	1911	5	PCT-US94-10166-5
13	42	47.2	1911	5	PCT-US94-10166-5
14	40	44.9	187	4	US-09-095-855-203
15	40	44.9	187	4	US-09-205-426-203
16	40	44.9	1097	4	US-09-252-991A-22579
17	39	43.8	16	3	US-08-931-220-2
18	39	43.8	16	5	PCT-US85-11723-2
19	39	43.8	16	5	PCT-US96-05997-15
20	39	43.8	124	4	US-09-107-532A-4773
21	39	43.8	241	4	US-09-134-001C-5493
22	39	43.8	398	5	US-08-931-220-5
23	39	43.8	398	5	PCT-US95-11723-5
24	39	43.8	398	5	PCT-US96-05997-1
25	38	42.7	133	4	US-09-134-001C-4122
26	38	42.7	269	4	US-09-252-991A-21483
27	38	42.7	414	2	US-08-984-171-1

28	38	42.7	425	4	US-09-328-352-4837	Sequence 4837, Ap
29	38	42.7	720	1	US-07-731-157A-2	Sequence 2, Appl
30	38	42.7	720	2	US-08-541-780-2	Sequence 2, Appl
31	37	41.6	94	4	US-09-107-532A-6937	Sequence 6937, Ap
32	37	41.6	170	4	US-09-732-210-1676	Sequence 1676, Ap
33	37	41.6	173	1	US-08-062-472B-6	Sequence 6, Appl
34	37	41.6	373	3	US-09-039-198A-14	Sequence 14, Appl
35	37	41.6	373	3	US-09-039-198A-15	Sequence 15, Appl
36	37	41.6	373	4	US-08-877-599-14	Sequence 14, Appl
37	37	41.6	373	4	US-08-877-599-15	Sequence 15, Appl
38	37	41.6	373	4	US-09-267-574-14	Sequence 14, Appl
39	37	41.6	373	4	US-09-267-574-15	Sequence 15, Appl
40	37	41.6	383	4	US-09-134-001C-3701	Sequence 3701, Ap
41	37	41.6	387	2	US-08-486-839-6	Sequence 6, Appl
42	37	41.6	387	2	US-09-151-011-6	Sequence 6, Appl
43	37	41.6	387	4	US-09-343-623-6	Sequence 6, Appl
44	37	41.6	415	4	US-09-554-999-2	Sequence 2, Appl
45	37	41.6	466	2	US-08-486-839-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
Prior Application Number:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERTHEMA INFECTION)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
US-08-856-841-13
Query Match 100.0%; Score 89; DB 3; Length 395;
Best Local Similarity 100.0%; Pred. No. 2,8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TYPEPKNKTOOYTOIOE 16
DB 252 TYPEPKNKTOOYTOIOE 267

APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 398
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 21:
US-08-856-841-21

RESULT 2
US-08-856-841-21
Sequence 21, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:

Query Match 100.0%; Score 89; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
DB 241 TYFPNKGTOQYTDQIE 256

RESULT 3

US-08-856-841-20
Sequence 20, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 415
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOMAS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 89; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
DB 252 TYFPNKGTOQYTDQIE 267

RESULT 4

US-08-856-841-22
Sequence 22, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDONS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 89; DB 3; Length 543;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIE 16
DB 400 TYFPNKGTOOYTDOIE 415

RESULT 5
US-09-438-268-4
Sequence 4, Application US/09438268
Patent No. 6491907
GENERAL INFORMATION:
APPLICANT: Rabinowitz, Joseph E.
APPLICANT: Samulecki, Richard J
TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 5470-186
CURRENT APPLICATION NUMBER: US/09/438,268
CURRENT FILING DATE: 1999-11-10
EARLIER APPLICATION NUMBER: 60/107,840
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/123,651
EARLIER FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 59
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 756
TYPE: PRT
ORGANISM: Virus
US-09-438-268-4

Query Match 100.0%; Score 89; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 5.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIE 16
DB 400 TYFPNKGTOOYTDOIE 415

DB 593 TYFPNKGTOOYTDOIE 608

RESULT 6
US-09-345-236B-135
Sequence 135, Application US/09345236B
Patent No. 6521454
GENERAL INFORMATION:
APPLICANT: Becnel, James J.
APPLICANT: Tukuo, Fukuda
APPLICANT: Moser, Bettina
APPLICANT: Cockburn, Andrew
APPLICANT: White, Susan E.
TITLE OF INVENTION: No. 6521454e1 Baculoviruses, Insecticidal
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates
FILE REFERENCE: 21042.0004
CURRENT APPLICATION NUMBER: US/09/345,236B
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 148
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 135
LENGTH: 119
TYPE: PRT
ORGANISM: mosquito baculovirus
US-09-345-236B-135

Query Match 49.4%; Score 44; DB 4; Length 119;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIE 16
DB 88 TCFPSTGTSEYTSWIK 103

RESULT 7
US-09-345-236B-108
Sequence 108, Application US/09345236B
Patent No. 6521454
GENERAL INFORMATION:
APPLICANT: Becnel, James J.
APPLICANT: Tukuo, Fukuda
APPLICANT: Moser, Bettina
APPLICANT: Cockburn, Andrew
APPLICANT: White, Susan E.
APPLICANT: Undeen, Albert H.
TITLE OF INVENTION: No. 6521454e1 Baculoviruses, Insecticidal
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates
FILE REFERENCE: 21042.0004
CURRENT APPLICATION NUMBER: US/09/345,236B
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 148
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 108
LENGTH: 235
TYPE: PRT
ORGANISM: mosquito baculovirus
US-09-345-236B-108

Query Match 49.4%; Score 44; DB 4; Length 235;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOIE 16
DB 204 TCFPSTGTSEYTSWIK 219

RESULT 8
US-08-447-464-3
Sequence 3, Application US/08447464
Patent No. 5840842

GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Yan, Hai
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/130,570
FILING DATE: 01-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-447-464-3

Query Match 47.2%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPKNKGTQY 11
|:|:|:|:|
1057 YPKNKGTETY 1066

RESULT 9
US-08-716-679-3
Sequence 3, Application US/08716679
Patent No. 5846800
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Yan, Hai
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/716,679
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/130,570
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-716-679-3

Query Match 47.2%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPKNKGTQY 11
|:|:|:|:|
Db 1057 YPKNKGTETY 1066

RESULT 10
US-08-348-006B-5
Sequence 5, Application US/08348006B
Patent No. 5658756
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND
STREET: 126 E. LINCOLN AVE., P.O. BOX 2060
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,006B
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J. MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 189921A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-825A-5

Query Match 47.2%; Score 42; DB 1; Length 1911;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
|:|:|:|:|:
Db 1467 YWPNRGTEY 1476

RESULT 11
US-08-800-825A-5
Sequence 5, Application US/08800825A
Patent No. 5865397

GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,825A
FILING DATE: 14-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J. MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18992DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3905
TELEFAX: 732-594-4720

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-825A-5

Query Match 47.2%; Score 42; DB 2; Length 1911;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
|:|:|:|:|:
Db 1467 YWPNRGTEY 1476

RESULT 12
US-09-158-657-5
Sequence 5, Application US/09158657
Patent No. 6214564

GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE

TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/158,657
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/800,825
FILING DATE: 14-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J. MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 18992DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3905
TELEFAX: 732-594-4720

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-158-657-5

Query Match 47.2%; Score 42; DB 3; Length 1911;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
|:|:|:|:|:
Db 1467 YWPNRGTEY 1476

RESULT 13
PCT-US94-10166-5
Sequence 5, Application PC/TUS9410166
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOHN W. WALLER III
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10166
FILING DATE: 09-SEPT-1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WALLEN, JOHN W III
REGISTRATION NUMBER: 35403
REFERENCE/DOCKET NUMBER: 18992
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-10166-5

Query Match
Best Local Similarity 47.2%; Score 42; DB 5; Length 1911;
Est Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFNGKGTQY 11
DB 1467 YFNGKGTETY 1476

RESULT 14
US-09-995-855-203
Sequence 203, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleach, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000,1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 203:

SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-995-855-203

Query Match
Best Local Similarity 44.9%; Score 40; DB 3; Length 187;
Est Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIE 16
DB 161 DKSTHQYTNQID 172

RESULT 15
US-09-205-426-203
Sequence 203, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
FILE REFERENCE: 11000,1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
US-09-205-426-203

Query Match
Best Local Similarity 44.9%; Score 40; DB 4; Length 187;
Est Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTQYTDQIE 16
DB 161 DKSTHQYTNQID 172

RESULT 16
US-09-252-991A-22579
Sequence 22579, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 22579
LENGTH: 1097
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22579

Query Match 44.9%; Score 40; DB 4; Length 1097;
Best Local Similarity 46.7%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 YPNKGTQOYDTQI 15
DB 769 TEPFRGQQQQGQEV 783

RESULT 17
US-08-931-220-2
Sequence 2, Application US/08931220
Patent No. 6030835

GENERAL INFORMATION:

APPLICANT: Musser M.D., James M.
TITLE OF INVENTION: Methods and Compositions for Identifying
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSER: WEIL, GOTSCHAL & MANGES
STREET: 2882 Sand Hill Road, Suite 280
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-7022

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,220
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
APPLICATION NUMBER: US 08/160,965
FILING DATE: 02-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: BAYL-004/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
POSITION IN GENOME:
MAP POSITION: 203
US-08-931-220-2

Query Match 43.8%; Score 39; DB 3; Length 16;
Best Local Similarity 60.0%; Pred. No. 2.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 FPNKGTQOYT 12
DB 4 YPNKGLKDYT 13

RESULT 18
PCT-US95-11723-2
Sequence 2, Application PC/TUS9511723
GENERAL INFORMATION:

APPLICANT: Musser M.D., James M.
TITLE OF INVENTION: Methods and Compositions for
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSER: LAW OFFICES OF BARBARA RAE-VENTER
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11723
FILING DATE: 14-SEP-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/03US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6205
TELEFAX: (415) 424-8760

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
POSITION IN GENOME:
MAP POSITION: 203
PCT-US95-11723-2

Query Match 43.8%; Score 39; DB 5; Length 16;
Best Local Similarity 60.0%; Pred. No. 2.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 FPNKGTQOYT 12
DB 4 YPNKGLKDYT 13

RESULT 19
PCT-US96-05997-15

Sequence 15, Application PC/TUS9605997

GENERAL INFORMATION:

APPLICANT: Musser M.D., James M.
TITLE OF INVENTION: Methods and Compositions for
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSER: ANANTHA SWAMY, H. N.
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

TITLE OF INVENTION: Use of extracellular cysteine protease
TITLE OF INVENTION: to inhibit cell proliferation
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEIL, GOSHAL & MANGES
STREET: 2882 Sand Hill Road, Suite 280
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-7022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05997
FILING DATE: 01-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/160,965
FILING DATE: 02-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/0205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
FAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
POSITION IN GENOME:
MAP POSITION: 203
PCT-US96-05997-15
Query Match 43.8%; Score 39; DB 5; Length 16;
Best Local Similarity 60.0%; Pred. No. 2.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 PPNKGTQOYT 12
DB 4 YPNKGKQOYT 13
RESULT 20
US-09-107-532A-4773
Sequence 4773, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-8277
FAX: (781) 893-5007
INFORMATION FOR SEQ ID NO: 4773:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...124
SEQUENCE DESCRIPTION: SEQ ID NO: 4773:
US-09-107-532A-4773
Query Match 43.8%; Score 39; DB 4; Length 124;
Best Local Similarity 58.3%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 3 PPNKGTQOYTQ 14
DB 64 PPNKGTQOYTQ 75
RESULT 21
US-09-134-001C-5493
Sequence 5493, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5493
LENGTH: 241
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5493
Query Match 43.8%; Score 39; DB 4; Length 241;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 YPNKGTQOYT 12
DB 223 YPNKGTQOYT 233

RESULT 22
US-08-931-220-5
Sequence 5, Application US/08931220
Patent No. 6030835
GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.
APPLICANT: Kapur M.D., Vivek
TITLE OF INVENTION: Methods and Compositions for Identifying
TITLE OF INVENTION: Group A Streptococcus
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEIL, GOTSAL & MANGES
STREET: 2882 Sand Hill Road, Suite 280
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-7022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,220
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
APPLICATION NUMBER: US 08/160,965
FILING DATE: 02-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: MGAS 1719
IMMEDIATE SOURCE:
CLONE: SPB7 (cysteine protease)
PUBLICATION INFORMATION:
AUTHORS: Kapur, V.
AUTHORS: Topouzis, S.
AUTHORS: Majesky, M. W.
AUTHORS: Li, L.-L.
AUTHORS: Hamrick, M. R.
AUTHORS: Hamill, R. J.
AUTHORS: Patil, J. M.
AUTHORS: Musser, J. M.
TITLE: A conserved Streptococcus pyogenes
TITLE: extracellular cysteine protease cleaves human
TITLE: fibronectin and degrades vitronectin
JOURNAL: Microb. Pathog.
VOLUME: 15
PAGES: 327-346
DATE: 1993
US-08-931-220-5

Query Match 43.8%; Score 39; DB 3; Length 398;
Best Local Similarity 60.0%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 FPKGTQOYT 12
DB 206 YPNKGKIDYT 215
RESULT 23
PCT-US95-11723-5
Sequence 5, Application PC/TUS9511723
GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.
APPLICANT: Kapur M.D., Vivek
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Identifying Streptococcus
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAW OFFICES OF BARBARA RAE-VENTER
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11723
FILING DATE: 14-SEP-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/03US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6205
TELEFAX: (415) 424-8760
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: MGAS 1719
IMMEDIATE SOURCE:
CLONE: SPB7 (cysteine protease)
PUBLICATION INFORMATION:
AUTHORS: Kapur, V.
AUTHORS: Topouzis, S.
AUTHORS: Majesky, M. W.
AUTHORS: Li, L.-L.
AUTHORS: Hamrick, M. R.
AUTHORS: Hamill, R. J.
AUTHORS: Patil, J. M.
AUTHORS: Musser, J. M.
TITLE: A conserved Streptococcus pyogenes
TITLE: extracellular cysteine protease cleaves human
TITLE: fibronectin and degrades vitronectin
JOURNAL: Microb. Pathog.
VOLUME: 15

PAGES: 327-346
DATE: 1993
PCT-US95-11723-5

Query Match 43.8%; Score 39; DB 5; Length 398;
Best Local Similarity 60.0%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPNKGTQOYT 12
:||||:|
DB 206 YPNKGLKDYT 215

RESULT 24
PCT-US96-05997-1

; Sequence 1, Application PC/TUS9605997
; GENERAL INFORMATION:
; APPLICANT: Musser M.D., James M.
; APPLICANT: Kapur M.D., Vivek
; APPLICANT: Ananthaswamy, H. N.
; APPLICANT: Fernandez, A.
; TITLE OF INVENTION: Use of extracellular cysteine protease
; TITLE OF INVENTION: to inhibit cell proliferation
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: WEIL, GOTSHAL & MANGES
; STREET: 2882 Sand Hill Road, Suite 280
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025-7022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05997
; FILING DATE: 01-MAY-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,542
; FILING DATE: 14-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/160,965
; FILING DATE: 02-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter Ph.D., Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: BAYL-004/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 926-6200
; TELEFAX: (415) 854-3713
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: MGAS 1719
; IMMEDIATE SOURCE:
; CLONE: speB7
; PCT-US96-05997-1

Query Match 43.8%; Score 39; DB 5; Length 398;
Best Local Similarity 60.0%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPNKGTQOYT 12
:||||:|
DB 206 YPNKGLKDYT 215

RESULT 25
US-09-134-001C-4122

; Sequence 4122, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4122
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4122

Query Match 42.7%; Score 38; DB 4; Length 133;
Best Local Similarity 43.8%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TYFPNKGTYQOYTDOIE 16
:|:|:|:|:|:|
DB 114 TYVYNGKGYDITSKIK 129

RESULT 26

US-09-252-991A-21483
; Sequence 21483, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21483
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21483

Query Match 42.7%; Score 38; DB 4; Length 269;
Best Local Similarity 50.0%; Pred. No. 79;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTYQOYTDO 14
:|:|:|:|:|:|
DB 102 TYIDAEGRVYTTDQ 115

RESULT 27

US-08-984-171-1
; Sequence 1, Application US/08984171
; Patent No. 5952177
; GENERAL INFORMATION:

```

APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Au-Yang, Janice
TITLE OF INVENTION: HUMAN CYTOSOLIC ISOCITRATE
DEHYDROGENASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,171
FILING DATE: Filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 414 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTTUT03
CLONE: 1996789
US-08-984-171-1

Query Match          42.7%; Score 38; DB 2; Length 414;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

1 TYFPNKGTOQYT 12
|||:|||||
155 TYTPSDGTQKVT 166
DB

RESULT 28
US-09-328-352-4837
Sequence 4837, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4837
LENGTH: 425
TYPE: PRP
ORGANISM: Acinetobacter baumannii
US-09-328-352-4837

Query Match          42.7%; Score 38; DB 4; Length 425;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;

```

```

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYT 12
|||:|||||
DB 160 TYTPSGTNKWT 171
DB

RESULT 29
US-07-731-157A-2
Sequence 2, Application US/07731157A
Patent No. 5457032
GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misset, Onno
APPLICANT: Van der Laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: COOLEY GODWARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157A
FILING DATE: 19910509
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAB-VENTER PH. D., BARBARA
REGISTRATION NUMBER: 32,750
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-731-157A-2

Query Match          42.7%; Score 38; DB 1; Length 720;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDQIE 16
|||:|||||
DB 680 GTTHVSDQIE 689
DB

RESULT 30
US-08-541-780-2
Sequence 2, Application US/08541780
Patent No. 5935831
GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misset, Onno
APPLICANT: Van der Laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,780
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157
FILING DATE:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER PH.D., BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GBRO-027/0005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-541-780-2

Query Match 42.7%; Score 38; DB 2; Length 720;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GTQGYTDQIE 16
|||:||||
Db 680 GTTHYSQDIE 689

RESULT 31
US-09-107-532A-6937
Sequence 6937, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6937:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...94
SEQUENCE DESCRIPTION: SEQ ID NO: 6937:
US-09-107-532A-6937

Query Match 41.6%; Score 37; DB 4; Length 94;
Best Local Similarity 41.7%; Pred. No. 35;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 PNKGTQGYTDQI 15
|||:||||
Db 16 PKNKTEYVSRL 27

RESULT 32
US-09-732-210-1676
Sequence 1676, Application US/097322210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mitranck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIORITY APPLICATION NUMBER: US 60/169,513
PRIORITY FILING DATE: 1999-12-07
PRIORITY APPLICATION NUMBER: US 60/169,340
PRIORITY FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1676
LENGTH: 170
TYPE: PRT
ORGANISM: Streptomyces coelicolor
US-09-732-210-1676

Query Match 41.6%; Score 37; DB 4; Length 170;
Best Local Similarity 53.8%; Pred. No. 69;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPNKGTOGYTDQ 14
|||:||||
Db 77 YFPNKHQOEYNE 89

RESULT 33
US-08-062-472B-6
Sequence 6, Application US/08062472B
Patent No. 5695954
GENERAL INFORMATION:
APPLICANT: Sherwood, Nancy G M

APPLICANT: Parker, David B
APPLICANT: McRory, John E
APPLICANT: Leischeld, David W
TITLE OF INVENTION: DNA ENCODING TWO FISH NEUROPEPTIDES
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: KLAUQUIST, SPARKMAN, CAMPBELL, LEIGH &
ADDRESSEE: WHINSTON, LLP
STREET: ONE WORLD TRADE CENTER, SUITE 1600, 121 S.W.
STREET: SALMON STREET
CITY: PORTLAND
STATE: OREGON
COUNTRY: USA
ZIP: 97204-2988
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/062,472B
FILING DATE: 14-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: POLLEY, RICHARD J
REGISTRATION NUMBER: 28107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-062-472B-6

Query Match 41.6%; Score 37; DB 1; Length 173;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 YFENKGTQOYTD 13
DB 73 YPEKGTERRHAD 84

SULT 34
US-09-039-198A-14
Sequence 14, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 41.6%; Score 37; DB 3; Length 373;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTOOYTDQI 15
DB 79 NFGTOKFTDMV 89

RESULT 35
US-09-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 41.6%; Score 37; DB 3; Length 373;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTOOYTDQI 15
DB 79 NFGTOKFTDMV 89

RESULT 36
US-08-877-599-14
; Sequence 14, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-877-599-14

Query Match 41.6%; Score 37; DB 4; Length 373;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

5 NKGTQYTDQI 15
| | | | |
79 NFGTKFTDMV 89

Db

RESULT 37
US-08-877-599-15
; Sequence 15, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-877-599-15

Query Match 41.6%; Score 37; DB 4; Length 373;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

5 NKGTQYTDQI 15
| | | | |
79 NFGTKFTDMV 89

Db

RESULT 38
US-09-267-574-14
; Sequence 14, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; EARLIER FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 14
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-267-574-14

Query Match 41.6%; Score 37; DB 4; Length 373;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

5 NKGTQYTDQI 15
| | | | |
79 NFGTKFTDMV 89

Db

RESULT 39
US-09-267-574-15
; Sequence 15, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407

;
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 373
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-267-574-15

Query Match 41.6%; Score 37; DB 4; Length 373;
Best local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTQOYTDQI 15
DB 79 NFGTQKFTDMV 89

RESULT 40
US-09-134-001C-3701
; Sequence 3701, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3701
; LENGTH: 383
; TYPE: PRF
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3701

Query Match 41.6%; Score 37; DB 4; Length 383;
Best local Similarity 37.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 12; Gaps 1;

2 YPPNKGSTQO-----YTDQIE 16
212 YGPOKGAQOKMIPKLDASALRHVHDKIE 238

Search completed: August 20, 2003, 09:34:27
Job time : 18.6495 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:30:58 ; Search time 64.9897 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-3

Perfect score: 89

Sequence: 1 TYFPNKGTOQYTDQIE 16

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

1 number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	756	US-10-205-942-4	Sequence 4, Appli
2	42	47.2	437	US-10-156-761-10200	Sequence 10200, A
3	42	47.2	1502	US-09-808-602-54	Sequence 54, Appli
4	42	47.2	1502	US-09-800-198-44	Sequence 44, Appli
5	42	47.2	1948	US-09-808-602-55	Sequence 55, Appli
6	42	47.2	1948	US-09-800-198-45	Sequence 45, Appli
7	41	46.1	283	US-09-815-242-13476	Sequence 13476, A
8	40	44.9	187	US-10-205-979-49	Sequence 49, Appli
9	40	44.9	187	US-10-051-643-203	Sequence 203, Appli
10	40	44.9	869	US-10-190-435-4	Sequence 4, Appli
11	40	44.9	1244	US-09-815-915-8	Sequence 8, Appli
12	39	43.8	248	US-10-003-784A-25	Sequence 25, Appli
13	39	43.8	398	US-10-002-784A-24	Sequence 24, Appli
14	39	43.8	468	US-10-002-784A-27	Sequence 27, Appli
15	39	43.8	605	US-10-238-075-691	Sequence 691, App

16	39	43.8	713	US-10-116-519-8	Sequence 8, Appli
17	38	42.7	405	US-10-029-180-78	Sequence 78, Appli
18	38	42.7	426	US-09-925-300-1502	Sequence 1502, App
19	38	42.7	434	US-10-029-180-112	Sequence 112, App
20	37.5	42.1	543	US-10-156-761-112	Sequence 13485, A
21	37	41.6	34	US-10-218-102-226	Sequence 126, App
22	37	41.6	173	US-10-156-761-12492	Sequence 12492, A
23	37	41.6	233	US-10-188-586A-7	Sequence 7, Appli
24	37	41.6	251	US-10-156-761-9227	Sequence 9227, Ap
25	37	41.6	373	US-10-161-547-14	Sequence 14, Appli
26	37	41.6	373	US-10-161-547-15	Sequence 15, Appli
27	37	41.6	414	US-10-108-605-87	Sequence 87, Appli
28	37	41.6	415	US-10-188-586A-2	Sequence 2, Appli
29	37	41.6	428	US-09-347-131-10	Sequence 10, Appli
30	37	41.6	428	US-09-563-728A-28	Sequence 28, Appli
31	37	41.6	433	US-09-859-888-4	Sequence 4, Appli
32	37	41.6	445	US-10-004-219B-10	Sequence 10, Appli
33	37	41.6	466	US-10-161-547-2	Sequence 2, Appli
34	37	41.6	472	US-09-745-763-36	Sequence 36, Appli
35	37	41.6	473	US-09-984-271-120	Sequence 120, App
36	37	41.6	478	US-09-859-888-2	Sequence 2, Appli
37	37	41.6	1076	US-10-116-949-6	Sequence 6, Appli
38	37	41.6	1145	US-10-116-949-2	Sequence 2, Appli
39	37	41.6	1145	US-10-116-949-4	Sequence 4, Appli
40	36	40.4	85	US-09-864-761-3996	Sequence 3996, A
41	36	40.4	92	US-10-106-698-6596	Sequence 6596, A
42	36	40.4	168	US-09-922-217-198	Sequence 198, App
43	36	40.4	168	US-09-833-263-198	Sequence 198, App
44	36	40.4	168	US-10-025-380-198	Sequence 198, App
45	36	40.4	254	US-09-815-242-5434	Sequence 5434, Ap

ALIGNMENTS

RESULT 1

US-10-205-942-4

Sequence 4, Application US//10205942

Publication No. US20030053990A1

GENERAL INFORMATION:

APPLICANT: University of No. US20030053990A1th Carolina-Chapel Hill

APPLICANT: Rabinowitz, Joseph E.

APPLICANT: Samulski, Richard J.

APPLICANT: Xiao, Weidong

TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME

FILE REFERENCE: 5470-186

CURRENT APPLICATION NUMBER: US/10/205,942

CURRENT FILING DATE: 2002-07-26

NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4

LENGTH: 756

TYPE: PRT

ORGANISM: Adeno-associated virus

FEATURE:

NAME/KEY: misc feature

LOCATION: (1) - (2271)

OTHER INFORMATION: B19/AAV chimeric capsid coding sequence

US-10-205-942-4

Query Match

Best Local Similarity 100.0%; Score 89; DB 15; Length 756;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16

DB 593 TYFPNKGTOQYTDQIE 608

RESULT 2

US-10-156-761-10200

Sequence 10200, Application US/10156761

Publication No. US20030119018A1

```
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 10200
/ LENGTH: 437
/ TYPE: PRT
/ ORGANISM: Streptomyces avermitilis
US-10-156-761-10200
```

```
Query Match          47.2%; Score 42; DB 15; Length 437;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      3 FPNKGTQOYTDQ 14
DB      280 FPNPNAQGTIDQ 291
```

```
RESULT 3
US-09-808-602-54
/ Sequence 54, Application US/09808602
/ Patent No. US20020155115A1
/ GENERAL INFORMATION:
/ APPLICANT: Vernet, Corine A
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Shimkets, Richard A
/ APPLICANT: Herrman, John L
/ APPLICANT: Majumder, Kumud
/ APPLICANT: Mishra, Vishnu
/ APPLICANT: Mezes, Peter S
/ APPLICANT: MacDougall, John
/ TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 15966-697 CIP
/ CURRENT APPLICATION NUMBER: US/09/808,602
/ CURRENT FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 09/800,198
/ PRIOR FILING DATE: 2001-03-05
/ PRIOR APPLICATION NUMBER: 60/186,596
/ PRIOR FILING DATE: 2000-03-03
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 54
/ LENGTH: 1502
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-808-602-54
```

```
Query Match          47.2%; Score 42; DB 10; Length 1502;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 YFPNKGTQOY 11
DB      1058 YWPNRGTEY 1067
```

```
RESULT 4
US-09-800-198-44
/ Sequence 44, Application US/09800198
```

```
/ Publication No. US20030087816A1
/ GENERAL INFORMATION:
/ APPLICANT: Vernet, Corine AM
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Shimkets, Richard A
/ APPLICANT: Herrmann, John L
/ APPLICANT: Majumder, Kumud
/ APPLICANT: Mishra, Vishnu
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Rastelli, Luca
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-697
/ CURRENT APPLICATION NUMBER: US/09/800,198
/ CURRENT FILING DATE: 2001-03-05
/ PRIOR APPLICATION NUMBER: 60/186,596
/ PRIOR FILING DATE: 2000-03-03
/ NUMBER OF SEQ ID NOS: 98
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 44
/ LENGTH: 1502
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-800-198-44
```

```
Query Match          47.2%; Score 42; DB 11; Length 1502;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 YFPNKGTQOY 11
DB      1058 YWPNRGTEY 1067
```

```
RESULT 5
US-09-808-602-55
/ Sequence 55, Application US/09808602
/ Patent No. US20020155115A1
/ GENERAL INFORMATION:
/ APPLICANT: Vernet, Corine A
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Shimkets, Richard A
/ APPLICANT: Herrman, John L
/ APPLICANT: Majumder, Kumud
/ APPLICANT: Mishra, Vishnu
/ APPLICANT: Mezes, Peter S
/ APPLICANT: MacDougall, John
/ TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 15966-697 CIP
/ CURRENT APPLICATION NUMBER: US/09/808,602
/ CURRENT FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 09/800,198
/ PRIOR FILING DATE: 2001-03-05
/ PRIOR APPLICATION NUMBER: 60/186,596
/ PRIOR FILING DATE: 2000-03-03
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 55
/ LENGTH: 1948
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-808-602-55
```

```
Query Match          47.2%; Score 42; DB 10; Length 1948;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 YFPNKGTQOY 11
DB      1504 YWPNRGTEY 1513
```

```
RESULT 6
US-09-800-198-45
```

Sequence 45, Application US/09800198
Publication No. US20030087816A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corinne AM
APPLICANT: Fernandes, Elma
APPLICANT: Shinkels, Richard A
APPLICANT: Hermann, John L
APPLICANT: Majumder, Kumud
APPLICANT: Mishra, Vishna
APPLICANT: Mezes, Peter S
APPLICANT: Raestelli, Luca
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-697
CURRENT FILING DATE: 2001-03-05
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 45
LENGTH: 1948
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-198-45

Query Match 47.2%; Score 42; DB 11; Length 1948;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 YFPNKGTOOY 11
|:|:|:|:|:
Db 1504 YFPNKGTOOY 1513

RESULT 7
US-09-815-242-13476
Sequence 13476, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl U.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13476
LENGTH: 283
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13476

Query Match 46.1%; Score 41; DB 9; Length 283;
Best Local Similarity 46.7%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 TYPNKGTOOYTDQI 15
|:|:|:|:|:|:
Db 205 TSVPNKSGTYTETV 219

RESULT 8
US-10-205-979-49
Sequence 49, Application US/10205979
Publication No. US20030147861A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
APPLICANT: Abernethy, Nevin
TITLE OF INVENTION: Compounds and Methods for the Modulation
FILE REFERENCE: 11000.1063U
CURRENT APPLICATION NUMBER: US/10/205,979
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/308,446
PRIOR FILING DATE: 2001-07-26
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(187)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-205-979-49

Query Match 44.9%; Score 40; DB 12; Length 187;
Best Local Similarity 58.3%; Pred. No. 41;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 NKGTQOYTDQIB 16
|:|:|:|:|:|:
Db 161 DKSTHQYTNQID 172

RESULT 9
US-10-051-643-203
Sequence 203, Application US/10051643
Publication No. US20020197265A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Methods and Compounds for the Treatment
FILE REFERENCE: System using Mycobacterium Vaccae
CURRENT APPLICATION NUMBER: US/10/051,643
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US09/156,181
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 08/996,624
PRIOR FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
US-10-051-643-203

Query Match 44.9%; Score 40; DB 14; Length 187;
Best Local Similarity 58.3%; Pred. No. 41;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTOOYTDOIE 16
DB 161 DKSTHOYTNOID 172

RESULT 10
US-10-190-435-4
; Sequence 4, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR WEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
FILE REFERENCE: P18133.003 / 2302-18133
CURRENT APPLICATION NUMBER: US/10/190,435
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 319
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4

LENGTH: 869

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: TVL_8_5
US-10-190-435-4

Query Match 44.9%; Score 40; DB 12; Length 869;
Best Local Similarity 54.5%; Pred. No. 2,3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TYPNKGTOOY 11
DB 399 TYPNKGTYKY 409

RESULT 11

US-09-815-915-8
; Sequence 8, Application US/09815915
; Patent No. US20020025931A1
; GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
FILE REFERENCE: 38155-20006.00
CURRENT APPLICATION NUMBER: US/09/815,915
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,846
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8

LENGTH: 1244

TYPE: PRT
ORGANISM: Homo sapiens

US-09-815-915-8

Query Match 44.9%; Score 40; DB 9; Length 1244;
Best Local Similarity 46.2%; Pred. No. 3,4e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 PKNKGTOOYTDOIE 16
DB 626 PTAASEQYTDRLR 638

RESULT 12

US-10-002-784A-25
; Sequence 25, Application US/10002784A
; Publication No. US20030036644A1
; GENERAL INFORMATION:
; /33

APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/23/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 25

LENGTH: 248

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: mutant streptococcal pyrogenic exotoxin-B
US-10-002-784A-25

Query Match 43.8%; Score 39; DB 15; Length 248;
Best Local Similarity 60.0%; Pred. No. 82;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPNKGTOOYT 12
DB 61 YPNKGLKDYT 70

RESULT 13

US-10-002-784A-24
; Sequence 24, Application US/10002784A
; Publication No. US20030036644A1
; GENERAL INFORMATION:

/33
APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/23/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 24

LENGTH: 398

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: mutant streptococcal pyrogenic exotoxin B prosegment
US-10-002-784A-24

Query Match 43.8%; Score 39; DB 15; Length 398;
Best Local Similarity 60.0%; Pred. No. 1,4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPNKGTOOYT 12
DB 206 YPNKGLKDYT 215

RESULT 14

US-10-002-784A-27
; Sequence 27, Application US/10002784A
; Publication No. US20030036644A1
; GENERAL INFORMATION:
; /33

APPLICANT: Ulrich, Robert G.

TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/23/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 27
LENGTH: 468
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: mutant SpeA-mutant SpeB fusion
US-10-002-784A-27

Query Match 43.8%; Score 39; DB 15; Length 468;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

3 FPNKGTQOYT 12
:||||:|
Db 281 YFNKGLKDYT 290

RESULT 15
US-10-238-075-691
Sequence 691, Application US/10238075
Publication No. US20030148324A1
GENERAL INFORMATION:
APPLICANT: I.N.S.E.R.M.
TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat
FILE REFERENCE: BLANDINE
CURRENT APPLICATION NUMBER: US/10/238,075
CURRENT FILING DATE: 2002-09-10
PRIOR APPLICATION NUMBER: 0003145
PRIOR FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 1576
SOFTWARE: Patent version 3.1
SEQ ID NO 691
LENGTH: 605
TYPE: PRT
ORGANISM: Escherichia coli
US-10-238-075-691

Query Match 43.8%; Score 39; DB 12; Length 605;
Best Local Similarity 46.7%; Pred. No. 2.2e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

2 YFPNKGTOQYTDQIE 16
:||||:|
Db 282 YFPNENSPYTKID 296

RESULT 16
US-10-116-519-8
Sequence 8, Application US/10116519
Publication No. US20030114373A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYPEPTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
FILE REFERENCE: D0124 NP
CURRENT APPLICATION NUMBER: US/10/116,519
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/281,253
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/288,768
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/296,180
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/300,620

PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 145
SOFTWARE: Patent version 3.0
SEQ ID NO 8
LENGTH: 713
TYPE: PRT
ORGANISM: Mus musculus
US-10-116-519-8

Query Match 43.8%; Score 39; DB 15; Length 713;
Best Local Similarity 33.3%; Pred. No. 2.7e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 14; Gaps 1;

QY 1 TYFPNK-----GTQOYTDQIE 16
:||||:|
Db 499 TFEFNKESGDFLLRPFSEKKGAGTQELDDQIQ 528

RESULT 17
US-10-029-180-78
Sequence 78, Application US/10029180
Publication No. US20020182708A1
GENERAL INFORMATION:
APPLICANT: Call, Brian M.
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin T.
APPLICANT: Milna, G. Todd
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeffrey C.
APPLICANT: Trueheart, Josh
APPLICANT: Zhang, Lixin
TITLE OF INVENTION: No. US20020182708A1 Regulators of Fungal Gene Expression
FILE REFERENCE: MIC-004
CURRENT APPLICATION NUMBER: US/10/029,180
CURRENT FILING DATE: 2001-12-22
PRIOR APPLICATION NUMBER: US 60/257,431
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 138
SOFTWARE: Fastseq for Windows Version 4.0
SEQ ID NO 78
LENGTH: 405
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fungal gene
US-10-029-180-78

Query Match 42.7%; Score 38; DB 14; Length 405;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 YFPNKGTOQYTD 13
:||||:|
Db 180 YFANKGDSFVD 191

RESULT 18
US-09-925-300-1502
Sequence 1502, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: Patent Ver. 2.0

SEQ ID NO 1502
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-300-1502

Query Match 42.7%; Score 38; DB 10; Length 426;
Best Local Similarity 58.3%; Pred. No. 2.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YFPNKGTOOQYT 12
DB 167 YTFPSDGTOKVT 178

RESULT 19

US-10-029-180-112
Sequence 112, Application US/10029180
Publication No. US20020182708A1
GENERAL INFORMATION:
APPLICANT: Holtzman, Brian M.
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin T.
APPLICANT: Milne, G. Todd
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeffrey C.
APPLICANT: Trueheart, Josh
APPLICANT: Zhang, Lixin
TITLE OF INVENTION: No. US20020182708A1e1 Regulators of Fungal Gene Expression
FILE REFERENCE: MIC-004
CURRENT APPLICATION NUMBER: US/10/029,180
CURRENT FILING DATE: 2001-12-22
PRIOR APPLICATION NUMBER: US 60/257,431
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 138
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 112
LENGTH: 434
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fungal gene
US-10-029-180-112

Query Match 42.7%; Score 38; DB 14; Length 434;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

2 YFPNKGTOOQYTD 13
DB 180 YFANKGADSFVD 191

RESULT 20

US-10-156-761-13485
Sequence 13485, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 13485
LENGTH: 543
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-13485

Query Match 42.1%; Score 37.5; DB 15; Length 543;
Best Local Similarity 40.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

QY 2 YFPNKGTOOQYTDQIE 16
DB 252 HEPHRC-QEMADRYE 265

RESULT 21

US-10-218-102-226
Sequence 226, Application US/10218102
Publication No. US20030130827A1
GENERAL INFORMATION:
APPLICANT: Bentzien, Joerg
APPLICANT: Dahiyat, Bassil I.
APPLICANT: Desjarlais, John R.
APPLICANT: Hayes, Robert J.
APPLICANT: Vielmetter, Jost
TITLE OF INVENTION: Protein Design Automation for Protein Libraries
FILE REFERENCE: A-67229-11/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/218,102
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: US 09/927,790
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US 60/311,545
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: US 60/324,899
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/351,937
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/352,103
PRIOR FILING DATE: 2002-01-25
NUMBER OF SEQ ID NOS: 432
SOFTWARE: PatentIn version 3.1
SEQ ID NO 226
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-218-102-226

Query Match 41.6%; Score 37; DB 16; Length 34;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPNKGTOOQYTDQI 15
DB 18 YFNKDTQOSTESM 31

RESULT 22

US-10-156-761-12492
Sequence 12492, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262

;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 12492
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Streptomyces avermitilis
US-10-156-761-12492

Query Match 41.6%; Score 37; DB 15; Length 173;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPNKGTOQYTDQ 14
||| |
80 YFPNKGTOQEVNE 92

RESULT 23
US-10-188-586A-7
; Sequence 7, Application US/10188586A
; Publication No. US20030045438A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Processes From Gram Positive Organisms
; FILE REFERENCE: GC389-US-D1
; CURRENT APPLICATION NUMBER: US/10/188,586A
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/554,999
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: PCT/US98/27040
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: GB 9727471.6
; PRIOR FILING DATE: 1997-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-188-586A-7

Query Match 41.6%; Score 37; DB 15; Length 233;
Best Local Similarity 58.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 NKGTOQYTDQ 16
||| |
206 NKGKXPTDQ 217

RESULT 24
US-10-156-761-9227
; Sequence 9227, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30

;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 9227
;; LENGTH: 251
;; TYPE: PRT
;; ORGANISM: Streptomyces avermitilis
US-10-156-761-9227

Query Match 41.6%; Score 37; DB 15; Length 251;
Best Local Similarity 37.5%; Pred. No. 1.8e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 YFPNKGTOQYTDQ 16
||| |
Db 196 TLVPSKDRKRYKDRVE 211

RESULT 25
US-10-161-547-14
; Sequence 14, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-14

Query Match 41.6%; Score 37; DB 12; Length 373;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTOQYTDQ 15
||| |
Db 79 NFGTQKFTDMV 89

RESULT 26
US-10-161-547-15
; Sequence 15, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-15

Query Match 41.6%; Score 37; DB 12; Length 373;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 NKGTOOYTDQI 15
DB 79 NKGTOOYTDQI 89

RESULT 27

US-10-108-605-87
; Sequence 87, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stem, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-87

Query Match 41.6%; Score 37; DB 14; Length 414;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 PNRKTOOYTDQI 15
DB 225 PNRKTEYFKQV 236

RESULT 28

US-10-188-586A-2
; Sequence 2, Application US/10188586A
; Publication No. US20030045438A1
; GENERAL INFORMATION:
; APPLICANT: Beceall, David A.
; TITLE OF INVENTION: Proteases From Gram Positive Organisms
; FILE REFERENCE: GC389-US-DI
; CURRENT APPLICATION NUMBER: US/10/188,586A
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/554,999
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: PCT/US98/27040
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: GB 9727471.6
; PRIOR FILING DATE: 1997-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-188-586A-2

Query Match 41.6%; Score 37; DB 15; Length 415;
Best Local Similarity 58.3%; Pred. No. 3.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 NKGTOOYTDQI 16

DB 208 NGKKXPYTDQPE 219

RESULT 29

US-09-347-331-10
; Sequence 10, Application US/09347331
; Patent No. US20010010909A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Miao, Gou Han
; APPLICANT: Odell, Joan
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Chromatin Associated Proteins
; FILE REFERENCE: BB-1118-C
; CURRENT APPLICATION NUMBER: US/09/347,331
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 06/092,841
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-347-331-10

Query Match 41.6%; Score 37; DB 9; Length 428;
Best Local Similarity 46.7%; Pred. No. 3.3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQI 15
DB 344 TRIENQNSROYTDQI 358

RESULT 30

US-09-563-728A-28
; Sequence 28, Application US/09563728A
; Publication No. US20030078216A1
; GENERAL INFORMATION:
; APPLICANT: MacLeod, Alan R
; APPLICANT: Li, Zoumei
; APPLICANT: Besterman, Jeffrey M
; TITLE OF INVENTION: Inhibition of Histone Deacetylase
; FILE REFERENCE: 106101.229
; CURRENT APPLICATION NUMBER: US/09/563,728A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,287
; PRIOR FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-563-728A-28

Query Match 41.6%; Score 37; DB 11; Length 428;
Best Local Similarity 46.7%; Pred. No. 3.3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQI 15
DB 344 TRIENQNSROYTDQI 358

RESULT 31

US-09-859-888-4
; Sequence 4, Application US/09859888
; Patent No. US20020173459A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al

```

; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001239
; CURRENT APPLICATION NUMBER: US/09/859,888
; CURRENT FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Human
US-09-859-888-4

Query Match          41.6%; Score 37; DB 10; Length 433;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      2 YFPNKGTOQYTD 13
        |||:||||
        375 YFTNTSSYEYTD 386

RESULT 32
US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boel, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(445)
; OTHER INFORMATION: /note="Human chitotriosidase"
10-004-219B-10

Query Match          41.6%; Score 37; DB 15; Length 445;
Best Local Similarity 54.5%; Pred. No. 3.4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY      5 NKGTOQYTDQI 15
        |||:||||
        79 NFGTKFTDMV 89
        DB

RESULT 33
US-10-161-547-2
; Sequence 2, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
```

```

; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-2

Query Match          41.6%; Score 37; DB 12; Length 466;
Best Local Similarity 54.5%; Pred. No. 3.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY      5 NKGTOQYTDQI 15
        |||:||||
        100 NFGTKFTDMV 110
        DB

RESULT 34
US-09-745-763-36
; Sequence 36, Application US/09745763
; Patent No. US20020065394A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 219
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/745,763
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-745-763-36

Query Match          41.6%; Score 37; DB 9; Length 472;
Best Local Similarity 42.9%; Pred. No. 3.7e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY      2 YFPNKGTOQYTDQI 15
        |||:||||
        213 YSPHTGTOEYDQV 226
        DB
```

```
RESULT 35
; Sequence 120, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 120
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (473)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-120

Query Match          41.6%; Score 37; DB 11; Length 473;
Best Local Similarity 42.9%; Pred. No. 3.7e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 YFPNKGTOOYTDQI 15
DB      213 YSPHTGIGYQDGV 226

RESULT 36
US-09-859-888-2
; Sequence 2, Application US/09859888
; Patent No. US20020173459A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001239
; CURRENT APPLICATION NUMBER: US/09/859,888
; CURRENT FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Human
US-09-859-888-2

Query Match          41.6%; Score 37; DB 10; Length 478;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 YFPNKGTOOYTD 13
DB      420 YFTNTSSYEYTD 431

RESULT 37
US-10-116-949-6
```

```
; Sequence 6, Application US/10116949
; Publication No. US20030044911A1
; GENERAL INFORMATION:
; APPLICANT: Lerman, Michael I.
; APPLICANT: Minna, John D.
; APPLICANT: Latif, Farida
; APPLICANT: Wei, Ming-Hui
; APPLICANT: Sekido, Yoshitaka
; APPLICANT: Gao, Boning
; APPLICANT: Duh, Fuh-Mei
; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
; FILE REFERENCE: NIH-05043
; CURRENT APPLICATION NUMBER: US/10/116,949
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/470,443
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/114,359
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-949-6

Query Match          41.6%; Score 37; DB 15; Length 1076;
Best Local Similarity 42.9%; Pred. No. 9.2e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      3 FPNKGTQOYTDQIE 16
DB      694 FPNKAEDMTENPE 707

RESULT 38
US-10-116-949-2
; Sequence 2, Application US/10116949
; Publication No. US20030044911A1
; GENERAL INFORMATION:
; APPLICANT: Lerman, Michael I.
; APPLICANT: Minna, John D.
; APPLICANT: Latif, Farida
; APPLICANT: Wei, Ming-Hui
; APPLICANT: Sekido, Yoshitaka
; APPLICANT: Gao, Boning
; APPLICANT: Duh, Fuh-Mei
; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
; FILE REFERENCE: NIH-05043
; CURRENT APPLICATION NUMBER: US/10/116,949
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/470,443
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/114,359
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-949-2

Query Match          41.6%; Score 37; DB 15; Length 1145;
Best Local Similarity 42.9%; Pred. No. 9.9e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      3 FPNKGTQOYTDQIE 16
DB      763 FPNKAEDMTENPE 776

RESULT 39
```

```

US-10-116-949-4
; Sequence 4, Application US/10116949
; Publication No. US2003004911A1
; GENERAL INFORMATION:
; APPLICANT: Lerman, Michael I.
; APPLICANT: Minna, John D.
; APPLICANT: Latif, Farida
; APPLICANT: Wei, Ming-Hui
; APPLICANT: Sekido, Yoshitaka
; APPLICANT: Gao, Boning
; APPLICANT: Duh, Puh-Mei
; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
; FILE REFERENCE: NIH-05043
; CURRENT APPLICATION NUMBER: US/10/116,949
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/470,443
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/114,359
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-949-4

Query Match          41.6%; Score 37; DB 15; Length 1145;
Best Local Similarity 42.9%; Pred. No. 9,9e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 3 FPNKGTQOYTDQIE 16
|||:|:|
Db 763 FPNKAAEDMTENPE 776

RESULT 40
US-09-864-761-3996
; Sequence 3996, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3996
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: MAP TO AC004254.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: EST_HUMAN HIT: BF131624.1, EVALUATE 5.90e+00
US-09-864-761-3996

Query Match          40.4%; Score 36; DB 9; Length 85;
Best Local Similarity 43.8%; Pred. No. 80;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Oy 1 TYFPNKGTQOYTDQIE 16
|||:|:|
Db 55 TSFPLGTRPTQAL 70

```

Search completed: August 20, 2003, 10:08:16
Job time : 64.9897 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 17.9794 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-3
Perfect score: 89
Sequence: 1 TYFPNKGTOQYTDQIE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	781	1 VCPV19	coat protein VP1 -
2	46	51.7	458	2 B90625	NADH dehydrogenase
3	46	51.7	714	2 S76082	hypothetical prote
4	45	50.6	299	2 AG3278	hypothetical prote
5	43	48.3	89	2 I68530	hemoglobin alpha c
6	43	48.3	459	2 T11333	NADH2 dehydrogenas
7	43	48.3	547	2 JC4519	heat-shock protein
8	43	48.3	548	2 C64076	chaperonin groEL -
9	42	47.2	88	2 A44929	protein-tyrosine-p
10	42	47.2	398	2 I56540	protein-tyrosine-p
11	42	47.2	458	2 B90623	NADH dehydrogenase
12	42	47.2	1262	1 B48758	protein-tyrosine-p
13	42	47.2	1496	1 A48758	protein-tyrosine-p
14	42	47.2	1501	2 I58148	protein-tyrosine-p
15	42	47.2	1863	2 S46217	protein-tyrosine-p
16	42	47.2	1907	2 S50893	protein-tyrosine-p
17	41	46.1	283	2 B98061	conserved hypotet
18	41	46.1	286	2 F84954	spermidine synthet
19	41	46.1	362	2 T27272	hypothetical prote
20	41	46.1	458	2 B90615	NADH dehydrogenase
21	41	46.1	459	2 T11202	NADH2 dehydrogenas
22	41	46.1	906	2 G75362	aconitate hydratase
23	41	46.1	1437	2 T31093	hypothetical protei
24	41	46.1	13055	2 T16580	hypothetical prote
25	40.5	45.5	349	2 B86667	D-alanine-D-alanin
26	40.5	45.5	399	2 C90317	hypothetical prote
27	40	44.9	250	2 T25582	hypothetical prote
28	40	44.9	259	2 F70032	ABC transporter (A
29	40	44.9	283	2 H95194	hypothetical prote

30	40	44.9	455	2 B90619	NADH dehydrogenase
31	40	44.9	459	2 T11084	NADH2 dehydrogenas
32	40	44.9	508	1 VGVNFR	spike glycoprotein
33	40	44.9	512	2 B90442	hypothetical prote
34	40	44.9	626	2 B82273	accessory coloniza
35	40	44.9	725	2 UC1300	endo-beta-1,4-gluc
36	40	44.9	1250	2 T00454	hypothetical prote
37	40	44.9	2475	2 S35307	polyprotein pp220
38	39	43.8	141	1 HAMSB	hemoglobin alpha-D
39	39	43.8	141	2 C24625	hemoglobin alpha-D
40	39	43.8	245	2 T10407	immediate early pr
41	39	43.8	306	2 F64381	dihydroorotate oxi
42	39	43.8	320	2 A24815	calpain (BC 3.4.22
43	39	43.8	326	1 JN0307	glycoprotein VP7 p
44	39	43.8	398	2 A37768	streptococcal pyro
45	39	43.8	420	2 T32191	hypothetical prote

ALIGNMENTS

RESULT 1
VCPV19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A:Note: host Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A24299
R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.
J. Virol. 58, 921-936, 1986
A>Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A:Reference number: A24299; MUID:86200451; PMID:3701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-references: EMBL:M13178; NID:G333375; PIDN:AAA66867.1; PID:G333377
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 100.0%; Score 89; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 7.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIE 16
Db 618 TYFPNKGTOQYTDQIE 633

RESULT 2
B90625
NADH dehydrogenase chain 4 ND4 [imported] - Casuarinus casuarinus mitochondrion
C:Species: mitochondrion Casuarinus casuarinus
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B90625
R:Haddarath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A>Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Molecule type: DNA
A:Residues: 1-458 <KUR>
A:Cross-references: GB:NC_002778; NID:G14141873; PIDN:NP_115359.1; GSPDB:GN00162
C:Genetics:
A:Gene: ND4
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

Query Match 51.7%; Score 46; DB 2; Length 458;
Best Local Similarity 58.8%; Pred. No. 7.5;
Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 1 TYFPNKGTQOQY--DOI 15
|||||:|:|:|
DB 45 TYFPNKGLTSMWTIDOI 61

RESULT 3

S76082
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.

A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 02-Aug-2002

C:Accession: S76082

R:Name: T. Sato, S.; Kozaki, H.; Tanaka, A.; Asamiu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
B.

A:Reference number: S74322; UID:97061201; PMID:8905231

A:Accession: S76082

A:Note: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-714 <KAN>

A:Cross-references: EMBL:D63999; GB:AB001339; NID:G1001396; PIDN:BA10060.1; PID:dt01071
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

C:Superfamily: Nostoc sp. hypothetical protein all2707; dnaJ amino-terminal homology
P:6-70/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 51.7%; Score 46; DB 2; Length 714;
Best Local Similarity 53.3%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 TYFPNKGTQOQYTDIE 16
|||:|:|:|:|
DB 383 TYFNPVQOQYLEOLE 397

QY 2 TYFPNKGTQOQYTDIE 16
|||:|:|:|:|
DB 383 TYFNPVQOQYLEOLE 397

RESULT 4

AG3278
hypothetical protein BMEI0212 [imported] - *Brucella melitensis* (strain 16M)

C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AG3278
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Lee, T.; Ivanova,
M.; Masur, M.; Goldmann, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
C. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*

A:Reference number: AD3252; PMID:11756688

A:Accession: AG3278

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-239 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51394.1; PID:gl7982098; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0212

A:Map position: 1

Query Match 50.6%; Score 45; DB 2; Length 299;
Best Local Similarity 57.1%; Pred. No. 7;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 TYFPNKGTQOQYTDI 15
|||:|:|:|:|
DB 162 TYFNRGTPTVDR 175

QY 2 TYFPNKGTQOQYTDI 15
|||:|:|:|:|
DB 162 TYFNRGTPTVDR 175

QY 2 TYFPNKGTQOQYTDI 15
|||:|:|:|:|
DB 162 TYFNRGTPTVDR 175

QY 2 TYFPNKGTQOQYTDI 15
|||:|:|:|:|
DB 162 TYFNRGTPTVDR 175

RESULT 5

168530
hemoglobin alpha chain (clone alpha-28) - rat (fragment)

C:Species: *Rattus norvegicus* (Norway rat)

C:Species: *Rattus norvegicus* (Norway rat)

C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 19-May-2000

C:Accession: 168530

R:Name: V.M.; Gu, Y.L.; Au, D.M.; Wong, W.M.; Ma, C.W.; Cheng, L.Y.

Hemoglobin 17, 363-371, 1993

A:Title: Two new rat alpha-globin sequences as identified by the conserved region PCR.

A:Reference number: 154239; UID:94042225; PMID:8226096

A:Accession: 168530

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-89 <RES>

A:Cross-references: GB:S66558; NID:9439804

C:Genetics:

A:introns: 32/2

C:Superfamily: globin; globin homology

C:Keywords: blood; oxygen carrier

P:59/Binding site: oxygen (His) (distal axial ligand) #status predicted

Query Match 48.3%; Score 43; DB 2; Length 89;
Best Local Similarity 43.8%; Pred. No. 4.1;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TYFPNKGTQOQYTDIE 16
|||:|:|:|:|
DB 42 TYFPHNTSGSDQVK 57

QY 1 TYFPNKGTQOQYTDIE 16
|||:|:|:|:|
DB 42 TYFPHNTSGSDQVK 57

RESULT 6

T11333

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - *Corvus frugilegus* mitochondrion

C:Species: mitochondrion *Corvus frugilegus*

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002

C:Accession: T11333

R:Harlid, A.; Arnason, U.

Proc. R. Soc. Lond. B Biol. Sci. 266, 305-309, 1999

A:Title: Analyses of mitochondrial DNA nest ratite birds within the Neognathae-support

A:Reference number: 217262

A:Accession: T11333

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-459 <HAR>

A:Cross-references: EMBL:Y18522; PIDN:CA177204.1

C:Genetics:

A:Genome: mitochondrion

A:Note: NADH4

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 48.3%; Score 43; DB 2; Length 459;
Best Local Similarity 52.9%; Pred. No. 24;

Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 1 TYFPNKGTQOQY--DOI 15
|||:|:|:|:|
DB 45 TYFPGKLTHTWTIDOI 61

QY 1 TYFPNKGTQOQY--DOI 15
|||:|:|:|:|
DB 45 TYFPGKLTHTWTIDOI 61

RESULT 7

JC4519

heat-shock protein GroEL - *Pasteurella multocida*

C:Species: *Pasteurella multocida*

C:Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999

C:Accession: JC4519

R:Love, B.C.; Hansen, L.M.; Hirsch, D.C.

Gene 166, 179-180, 1995

A:Title: Cloning and sequence of the groEL heat-shock operon of *Pasteurella multocida*.

A:Reference number: JC4518; UID:96105224; PMID:8529887

A:Accession: JC4519

A:Molecule type: DNA

A:Residues: 1-547 <LOV>

A:Cross-references: GB:U01065; NID:G1144300; PIDN:AAA84916.1; PID:G1144302

A:Experimental source: serotype A:3

C:Genetics:

A:Gene: groEL

Db 2 YWPNRGTETY 11

C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 10-Sep-1999 #sequence_rev 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: B48758

R.Pan, M.G.; Rim, C.; Lu, K.P.; Florio, T.; Stork, P.J.S.
 J. Biol. Chem. 268, 19284-19291, 1993
 A>Title: Cloning and expression of two structurally distinct receptor-linked protein-tyr
 A:Reference number: A48758; MUID:93374907; PMID:8396131
 A:Accession: B48758
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1262 <PAN>
 A:Cross-references: GB:L19181
 A>Note: the authors translated the codon TGC for residue 27 as Gly, GAG for residue 79 a
 727 as Phe; residues 1237-1244 are shown after residue 1262, and, consequently, residues
 C:Superfamily: Leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; transm
 F:149-209/Domain: immunoglobulin homology <IMM1>
 F:246-300/Domain: immunoglobulin homology <IMM2>
 F:318-398/Domain: fibronectin type III repeat homology <FN3A>
 F:411-497/Domain: fibronectin type III repeat homology <FN3B>
 F:509-590/Domain: fibronectin type III repeat homology <FN3C>
 F:600-684/Domain: fibronectin type III repeat homology <FN3D>
 F:718-798/Domain: fibronectin type III repeat homology <FN3E>
 F:880-1496/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:1146/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1146/Binding site: substrate phosphate (Arg) #status predicted
 F:1435/Binding site: substrate phosphate (Arg) #status predicted

Query Match 47.2%; Score 42; DB 1; Length 1496;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPNNKGTQY 11
 |::|::|::|
 DB 1055 YWPNRGTEY 1064

RESULT 13
 A48758
 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor-linked form P1 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A48758
 R.Pan, M.G.; Rim, C.; Lu, K.P.; Florio, T.; Stork, P.J.S.
 J. Biol. Chem. 268, 19284-19291, 1993
 A>Title: Cloning and expression of two structurally distinct receptor-linked protein-tyr
 A:Reference number: A48758; MUID:93374907; PMID:8396131
 A:Accession: A48758
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1496 <PAN>
 A:Cross-references: GB:L19180
 A>Note: authors translated the codon TGC for residue 27 as Gly, GAG for residue 79 as Gl
 as Phe
 C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; transm
 F:149-209/Domain: immunoglobulin homology <IMM1>
 F:246-300/Domain: immunoglobulin homology <IMM2>
 F:318-405/Domain: fibronectin type III repeat homology <FN3A>
 F:411-504/Domain: fibronectin type III repeat homology <FN3B>
 F:509-599/Domain: fibronectin type III repeat homology <FN3C>
 F:600-684/Domain: fibronectin type III repeat homology <FN3D>
 F:718-798/Domain: fibronectin type III repeat homology <FN3E>
 F:880-1496/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:1146/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1146/Binding site: substrate phosphate (Arg) #status predicted
 F:1435/Binding site: Cys (phosphocysteine intermediate) #status predicted

DB 1055 YWPNRGTEY 1064

RESULT 14
 I58148
 protein-tyrosine-phosphatase (EC 3.1.3.48) 2B, splice form LAR - rat
 N:Alternate names: leukocyte common antigen-related phosphatase
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
 C:Accession: I58148; S46218
 R.Walton, K.M.; Martelli, K.J.; Kwak, S.P.; Dixon, J.E.; Lagent, B.L.
 Neuron 11, 387-400, 1993
 A>Title: A novel receptor-type protein tyrosine phosphatase is expressed during neurog
 A:Reference number: I58148; MUID:93357030; PMID:8352946
 A:Accession: I58148
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1501 <WAL>
 A:Cross-references: GB:L19933; NID:9310242; PID:AAA442309.1; PID:9310243
 A>Note: in Genbank entry RATTYRPHOS, release 113.0, the source is designated as Rattus
 R.Walton, W.R.; Hashimoto, N.; Ahmad, F.; Ding, W.; Goldstein, B.J.
 Biochem. J. 302, 39-47, 1994
 A>Title: Molecular cloning and expression of a unique receptor-like protein-tyrosine-ph
 A:Reference number: S46216; MUID:94347119; PMID:8068021
 A:Accession: S46218
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-1501 <ZHA>
 A:Cross-references: EMBL:L12329; NID:9294573; PID:NAC37657.1; PID:9294574
 C:Superfamily: Leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosi
 F:149-109/Domain: immunoglobulin homology <IMM1>
 F:149-209/Domain: immunoglobulin homology <IMM2>
 F:246-300/Domain: immunoglobulin homology <IMM3>
 F:313-506/Domain: fibronectin type III repeat homology <FN3A>
 F:382-1501/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:969-1190/Domain: protein-tyrosine-phosphatase homology <PTP2>
 F:1258-1481/Domain: protein-tyrosine-phosphatase homology <PTP3>
 F:1144/Active site: Cys (phosphocysteine intermediate) #link PTP1 #status predicted
 F:1144/Binding site: substrate phosphate (Arg) #link PTP1 #status predicted
 F:1433/Active site: Cys (phosphocysteine intermediate) #link PTP2 #status predicted
 F:1433/Binding site: substrate phosphate (Arg) #link PTP2 #status predicted

Query Match 47.2%; Score 42; DB 2; Length 1501;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YPNNKGTQY 11
 |::|::|::|
 DB 1057 YWPNRGTEY 1066

RESULT 15
 S46217
 protein-tyrosine-phosphatase (EC 3.1.3.48) type sigma precursor - rat
 N:Alternate names: leukocyte common antigen-related phosphatase
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 07-May-1995 #sequence_revision 03-Nov-1995 #text_change 23-Jul-1999.
 C:Accession: S46217; S51174; A45104
 R.Zhang, W.R.; Hashimoto, N.; Ahmad, F.; Ding, W.; Goldstein, B.J.
 Biochem. J. 302, 39-47, 1994
 A>Title: Molecular cloning and expression of a unique receptor-like protein-tyrosine-ph
 A:Reference number: S46216; MUID:94347119; PMID:8068021
 A:Accession: S46217
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1863 <ZHA>
 A:Cross-references: EMBL:L11587
 R:Goldstein, B.J.
 submitted to the EMBL Data Library, February 1993
 A:Reference number: S51174
 A:Accession: S51174

A.Molecule type: mRNA
 A.Residues: 1-1788, 'G', 1790-1863 <GOL>
 A.Cross-references: EMBL:L11587; NID:G205134; PIDN:AC37656.1; PID:G205135
 R.Yan, H.; Grossman, A.; Wand, H.; D'Eustachio, P.; Mossie, K.; Musacchio, J.M.; Silvent
 J. Biol. Chem. 268, 24880-24886, 1993
 A.Title: A novel receptor tyrosine phosphatase-sigma that is highly expressed in the ner
 A.Reference number: A49104; MUID:94043351; PMID:8227050
 A.Accession: A49104
 A.Status: preliminary; not compared with conceptual translation
 A.Molecule type: nucleic acid
 A.Residues: 1-596, 'R', 598-603, 'I', 967-1788, 'G', 1790-1863 <YAN>
 A.Experimental source: brain
 A.Note: sequence extracted from NCBI backbone (NCBIP:139669)
 A.Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy

C.Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
 F.1-26/Domains: signal sequence #status predicted <SIS>
 F.127-1863/Product: protein-tyrosine-phosphatase #status predicted <MAT>
 F.149-209/Domains: immunoglobulin homology <IMM1>
 F.46-300/Domains: immunoglobulin homology <IMM2>
 F.13-499/Domains: fibronectin type III repeat homology <FN3A>
 F.511-592/Domains: fibronectin type III repeat homology <FN3C>
 F.1244-1863/Domains: leukocyte common antigen cytosolic domain homology <LAC>
 F.1331-1552/Domains: protein-tyrosine-phosphatase homology <PTP1>
 F.1504/Active site: Cys (phosphocysteine intermediate) #status predicted
 F.1510/Binding site: substrate phosphate (Arg) #status predicted
 F.1795/Active site: Cys (phosphocysteine intermediate) #status predicted
 F.1801/Binding site: substrate phosphate (Arg) #status predicted

Query Match 47.2%; Score 42; DB 2; Length 1863;
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 DB 1419 YFPNKGTOOY 1428

RESULT 16

SS0893
 Protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type sigma precursor - mouse
 C.Species: Mus musculus (house mouse)
 C.Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000
 C.Accession: S50893; S40281
 R.Wagner, J.; Boerboom, D.; Tremblay, M.L.
 Eur. J. Biochem. 226, 773-782, 1994
 A.Title: Molecular cloning and tissue-specific RNA processing of a murine receptor-type
 A.Reference number: S50893; MUID:95112841; PMID:7529177
 A.Accession: S50893
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-1907 <WAG>
 A.Cross-references: EMBL:X82288; NID:G587483; PIDN:CA57732.1; PID:G587484
 R.Hendriks, W.; Brugman, C.; Zeeuwen, P.; Schepens, J.; Wieringa, B.
 submitted to the EMBL Data Library, June 1993
 A.Description: Assessment of the expression levels of murine protein-tyrosine phosphatase
 A.Reference number: S40280
 A.Accession: S40281
 A.Molecule type: mRNA
 A.Residues: 1441-1501, 'E', 1503-1546 <HEN>
 A.Cross-references: EMBL:Z23050; NID:G438137; PIDN:CA80585.1; PID:G438138
 A.Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy

C.Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane
 F.1149-209/Domains: immunoglobulin homology <IMM1>
 F.246-300/Domains: immunoglobulin homology <IMM2>
 F.413-506/Domains: fibronectin type III repeat homology <3FR>
 F.1288-1907/Domains: leukocyte common antigen cytosolic domain homology <LAC>
 F.1375-1596/Domains: protein-tyrosine-phosphatase homology <PTP1>
 F.1664-1887/Domains: protein-tyrosine-phosphatase homology <PTP2>
 F.1548/Active site: Cys (phosphocysteine intermediate) #status predicted
 F.1554/Binding site: substrate phosphate (Arg) #status predicted

F.1839/Active site: Cys (phosphocysteine intermediate) #status predicted
 F.1845/Binding site: substrate phosphate (Arg) #status predicted

Query Match 47.2%; Score 42; DB 2; Length 1907;
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 DB 1463 YFPNKGTOOY 1472

RESULT 17

E89061
 conserved hypothetical protein spr1518 [imported] - Streptococcus pneumoniae (strain R6)
 C.Species: Streptococcus pneumoniae
 C.Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C.Accession: E89061
 R.Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 Y. P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A.Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A.Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A.Reference number: A97872; MUID:21429245; PMID:11544234
 A.Accession: E89061
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-283 <KUN>
 A.Cross-references: GB:AE007317; PIDN:AAL00322.1; PID:G15459180; GSPDB:GN00174
 A.Gene: spr1518

Query Match 46.1%; Score 41; DB 2; Length 283;
 Best Local Similarity 46.7%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQI 15
 DB 205 TSVPNKGTOOYTDQI 219

RESULT 18

F84954
 spermidine synthase (EC 2.5.1.16) [imported] - Buchnera sp. (strain APS)
 C.Species: Buchnera sp.
 C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C.Accession: F84954
 R.Shigenobu, S.; Matanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A.Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A.Reference number: A84930; MUID:20445173; PMID:10993077
 A.Accession: F84954
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-286 <STO>
 A.Cross-references: GB:AP000398; GSPDB:GN00144
 A.Experimental source: strain APS
 C.Genetics:
 A.Gene: speE; BU209
 C.Superfamily: spermidine synthase
 C.Keywords: transferase

Query Match 46.1%; Score 41; DB 2; Length 286;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 YFPNKGTOOYTD 13
 DB 120 YFPNKGTOOYTD 131

RESULT 19

A:Cross-references: EMBL:AF090341; NID:g4894501; PID:g4894511; PIDN:AAD32534.1
C:Genetics:
A:Genome: mitochondrion
A:Gene code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 46.1%; Score 41; DB 2; Length 459;
Best Local Similarity 52.9%; Pred. No. 52;
Matches 9; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

Oy 1 TYPENKGTQOYT--DOI 15
||| |||
Db 45 TYYPGKGTLTPMTSIDOI 61

RESULT 22
G75362
acornate hydratase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: G75362
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radiation-resistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75362
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-906 <WHI>
A:Cross-references: GB:AE002013; GB:AE000513; NID:g6459484; PIDN:AAF11276.1; PID:g645949
C:Experimental source: strain R1
C:Genetics:
A:Map position: 1
C:Superfamily: iron-responsive element-binding protein

Query Match 46.1%; Score 41; DB 2; Length 906;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 2 YFPNKGTQOYTDQI 15
||| |||
Db 485 YLEPAGLQOYLDQI 498

RESULT 23
T31093
probable protein-tyrosine-phosphatase (EC 3.1.3.48) - medicinal leech
N:Alternate names: receptor tyrosine phosphatase
C:Species: Hirudo medicinalis (medicinal leech)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T31093
R:Gereshon, T.R.; Baker, M.; Nitabach, M.; Wu, P.; Macagno, E.R.
Submitted to the EMBL Data Library, December 1997
A:Description: Two receptor tyrosine phosphatases expressed by neurons and muscle cells d
A:Reference number: Z20976
A:Accession: T31093
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1437 <GER>
A:Cross-references: EMBL:AF017084; NID:g2695656; PID:g2695657; PIDN:AAB91461.1
C:Genetics:
A:Gene: LARI
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
CGY
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane

Query Match 46.1%; Score 41; DB 2; Length 1437;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YFPNKGTOOY 11
 |||||
 DB 986 YWPNKGTVEY 995

RESULT 24

hypothetical protein K07E12.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16580
 R:Fulton, L.

A:Submitted to the EMBL Data Library, May 1994
 A:Description: The sequence of *C. elegans* cosmid K07E12.
 A:Reference number: Z18540

A:Accession: T16580
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-13055 <FULL>

A:Cross-references: EMBL:U00054; NID:G485140; PID:G485141; PIDN:AA50715.1; CESP:K07E12.
 A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K07E12.1

A:Introns: 46/3; 85/2; 201/1; 278/2; 470/2; 817/1; 927/1; 960/3; 1265/1; 1322/1; 1478/3;
 ; 6014/3; 6159/3; 6665/2; 7266/3; 7895/3; 8669/3; 8726/3; 9803/3; 10937/3; 12234/2; 1229

Query Match 46.1%; Score 41; DB 2; Length 13055;

Best Local Similarity 61.5%; Pred. No. 1.9e+03;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 PNPNGTOOYTD 16
 |||||
 DB 1879 PNEWTIOYTDDE 1891

RESULT 25

E86667

D-alanine-D-alanine ligase (EC 6.3.2.4) [imported] - *Lactococcus lactis* subsp. *lactis* (E

C:Species: *Lactococcus lactis* subsp. *lactis*

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Jun-2002

C:Accession: E86667

R:Boletín, A.; Wincker, P.; Manger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* se

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: E86667
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-349 <STO>
 A:Cross-references: GB:AE005176; PID:G12723209; PIDN:AAK04439.1; GSPDB:GN00146

A:Experimental source: strain IL4403

C:Superfamily: D-alanine-D-alanine ligase
 C:Keywords: ligase

Query Match 45.5%; Score 40.5; DB 2; Length 349;

Best Local Similarity 47.4%; Pred. No. 47;

Matches 9; Conservative 3; Mismatches 2; Indels 5; Gaps 1;

QY 1 TYFPNKG-----TOOYTDO 14
 |||||
 DB 40 TYFIRKGGEFIRKTOEFTDK 58

RESULT 26

C90317

hypothetical protein SSO1584 [imported] - *Sulfolobus solfataricus*

C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: C90317

R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001

A:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: C90317

A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-499 <KDR>

A:Cross-references: GB:AE006641; NID:G13814814; PIDN:AAK41794.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO1584

Query Match 45.5%; Score 40.5; DB 2; Length 499;

Best Local Similarity 50.0%; Pred. No. 69;

Matches 8; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

RESULT 27

T25562

hypothetical protein C32B5.7 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000

C:Accession: T25562
 R:Schnee, P.

A:Submitted to the EMBL Data Library, December 1996
 A:Description: The sequence of *C. elegans* cosmid C32B5.
 A:Reference number: Z20054

A:Accession: T25562
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-250 <SCH>
 A:Cross-references: EMBL:U80843; PIDN:AA837963.1; GSPDB:GN00020; CESP:C32B5.7

A:Experimental source: strain Bristol N2; clone C32B5

C:Genetics:

A:Gene: CESP:C32B5.7

A:Map position: 2
 A:Introns: 23/2; 138/2; 197/1

C:Superfamily: papain
 Query Match 44.9%; Score 40; DB 2; Length 250;

Best Local Similarity 53.8%; Pred. No. 40;

Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 28

F70032

ABC transporter (ATP-binding protein) homolog yycR - *Bacillus subtilis*

C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001.

C:Accession: F70032
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bortner

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Gallier
 lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Holo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis,
 A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y. M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetalle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Seror
 akendul, M.; Tamakoshi, A.; Tanaka, T.; Terpilata, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: F70032
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-259 <KUN>
A:Cross-references: GB:299121; GB:AL009126; NID:g2635827; PIDN:CAB15475.1; PID:el186158;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yycR
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:25-221/Domain: ATP-binding cassette homology <ABC>
F:42-49/Region: nucleotide-binding motif A (P-loop)

Query Match 44.9%; Score 40; DB 2; Length 259;
Best Local Similarity 63.6%; Pred. No. 42;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 TYFPNKGTOOYTDOJ 11
|||:|||||
Db 12 TVSPNKDSQTYTETV 22

RESULT 29
H95194
hypothetical protein SPI674 [imported] - *Streptococcus pneumoniae* (strain TIGR4)
C:Species: *Streptococcus pneumoniae*
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95194
R:Retelid, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75753.1; PID:g14973166; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI674

Query Match 44.9%; Score 40; DB 2; Length 283;
Best Local Similarity 46.7%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDOJ 15
|||:|||||
Db 205 TVSPNKDSQTYTETV 219

RESULT 30
B90619
NADH dehydrogenase chain 4 ND4 [imported] - *Apteryx haastii* mitochondrion
C:Species: mitochondrion *Apteryx haastii*
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C:Accession: B90619
R:Hardrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: B90619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <KUR>
A:Cross-references: GB:NC_002782; NID:g14141943; PIDN:NP_115411.1; GSPDB:GN00159
C:Genetics:
A:Gene: ND4
A:Gene: mitochondrion

A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: mitochondrion

Query Match 44.9%; Score 40; DB 2; Length 455;
Best Local Similarity 58.8%; Pred. No. 76;
Matches 10; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 1 TYFPNKGTOOYT--DOJ 15
|||||:|||||
Db 45 TYFPYKTSQWTGIDQI 61

RESULT 31
T11084
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - *Falco peregrinus* mitochondrion
C:Species: mitochondrion *Falco peregrinus*
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11084
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242
A:Accession: T11084
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <MIN>
A:Cross-references: EMBL:AF090338; NID:g4894462; PID:g4894472; PIDN:AA02501.1
C:Genetics:
A:Gene: mitochondrion
A:Genetic code: SGC1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NMD; oxidative phosphorylation;

Query Match 44.9%; Score 40; DB 2; Length 455;
Best Local Similarity 52.9%; Pred. No. 77;
Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 1 TYFPNKGTOOYT--DOJ 15
|||||:|||||
Db 45 TYFPKTLTQWTGMDQI 61

RESULT 32
VGWVNR
spike glycoprotein G precursor - infectious hematopoietic necrosis virus (strain Round B
C:Species: infectious hematopoietic necrosis virus
A:Note: host *Oncorhynchus tshawytscha* (Chinook salmon)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Sep-1994
C:Accession: A29532
R:Koener, J.F.; Passavant, C.W.; Kurath, G.; Leong, J.
J. Virol. 61, 1342-1349, 1987
A:Title: Nucleotide sequence of a cDNA clone carrying the glycoprotein gene of infection
A:Reference number: A29532; MUID:87198856; PMID:3033264
A:Accession: A29532
A:Molecule type: mRNA
A:Residues: 1-508 <KOR>
A:Cross-references: GB:M16023
C:Genetics:
A:Gene: G
C:Superfamily: rhabdovirus spike glycoprotein G
C:Keywords: glycoprotein; spike protein; transmembrane protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-508/Product: spike glycoprotein G #status predicted <SGS>
F:469-485/Domain: transmembrane #status predicted <TMN>
F:56,400,401,438/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.9%; Score 40; DB 1; Length 508;
Best Local Similarity 46.7%; Pred. No. 86;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 TYFPNKGTOOYTDOIB 16
|||||:|||||

Db 144 YFPAPRCQWYDNDQ 158

RESULT 33

E90442
hypothetical protein SSO2678 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: E90442
R:She, O.; Sindh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Ameyez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrect, R.A.; Kagan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90442
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <KUR>
A:Cross-references: GB:AE006641; NID:g13815996; PIDN:AAK42796.1; GSPDB:GN00155
Genetics:
Gene: SSO2678

Query Match 44.9%; Score 40; DB 2; Length 512;
Best Local Similarity 58.3%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQ 12
||| |
Db 156 TLFPNKSSONYT 167

RESULT 34

E82273
accessory colonization factor AcfB VC0840 [imported] - Vibrio cholerae (strain N16961 se
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82273
R:Heideberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82273
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-626 <HEI>
A:Cross-references: GB:AE004169; GB:AE003852; NID:g9655289; PIDN:AAF94003.1; GSPDB:GN001
C:Genetics:
Experimental source: serogroup O1, strain N16961, biotype El Tor
A:Gene: VC0840
A:Map position: 1

Query Match 44.9%; Score 40; DB 2; Length 626;
Best Local Similarity 50.0%; Pred. No. 11e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQ 14
||| |
Db 23 TYFQMGHGTNEYIDR 36

RESULT 35

JCI300
endo-beta-1,4-glucanase (EC 3.2.1.-) CelCCG precursor - Clostridium cellulolyticum
C:Species: Clostridium cellulolyticum
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
C:Accession: JCI300
R:Bagnara-Tardif, C.; Gaudin, C.; Belaich, A.; Hoest, P.; Citard, T.; Belaich, J.P.
Gene 119, 17-18, 1992
A>Title: Sequence analysis of a gene cluster encoding cellulases from Clostridium cellul
A:Reference number: JCI299; MUID:93012971; PMID:1398087

A:Accession: JCI300
A:Molecule type: DNA
A:Residues: 1-725 <BAG>
A:Cross-references: GB:M67018; NID:g5597001; PIDN:AAA73868.1; PID:g551774
C:Genetics:
A:Gene: celCCG
C:Superfamily: Clostridium cellulase repeat homology
C:Keywords: glycosidase; hydrolase
F:1-35/Domains: signal sequence #status predicted <STG>
F:35-725/Product: endo-beta-1,4-glucanase CelCCG #status predicted <MAT>
F:664-687/Domains: Clostridium cellulase repeat homology <CCR2>
F:696-719/Domains: Clostridium cellulase repeat homology <CCR>
F:90,93,160,408,455/Active site: Asp, Asp, His, His, Glu #status predicted

Query Match 44.9%; Score 40; DB 2; Length 725;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQ 15
||| |
Db 271 SYFPWKGKQYTDII 285

RESULT 36

T00454
hypothetical protein T14N5.11 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 22-Oct-1999
C:Accession: T00454
R:Federpspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo,
J.; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z14152
A:Accession: T00454
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1250 <FPD>
A:Cross-references: EMBL:AC004260; NID:g3176694; PID:g3540206; GSPDB:GN00059; ATSD:T14N5
C:Genetics:
A:Gene: ATSP:T14N5.11
A:Map position: 1
A:introns: 206/3; 347/1; 384/1; 437/2; 523/2; 564/2; 590/2; 645/3; 660/3; 692/3;

Query Match 44.9%; Score 40; DB 2; Length 1250;
Best Local Similarity 52.6%; Pred. No. 2.3e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 3 FPNKGTQ-----QYTDQ 15
||| |
Db 236 FPGGCTQPIHGADQYLDQI 254

RESULT 37

S35307
polyprotein pp220 precursor - African swine fever virus
N:Contains: structural protein p14; structural protein p150; structural protein p34; str
C:Species: African swine fever virus, ASFV
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 24-Nov-1999
C:Accession: S35307; S39809
R:Simon-Mateo, C.; Andreu, G.; Vinnela, E.
EMBO J. 12, 2977-2987, 1993
A>Title: Polyprotein processing in African swine fever virus: a novel gene expression st
A:Reference number: S35307; MUID:93327788; PMID:8335009
A:Accession: S35307
A:Molecule type: DNA
A:Residues: 1-2475 <SIM>
A:Cross-references: EMBL:Z22777; NID:g394708; PIDN:CAA80455.1; PID:g394709
A:Accession: S39809
A:Molecule type: protein
A:Residues: 45-51/523-549/894-908 <S12>
C:Superfamily: African swine fever virus polyprotein pp220
C:Keywords: myristylation; polyprotein

F:45-522/Product: structural protein p34 #status predicted <MAT1>
F:523-893/Product: structural protein p37 #status predicted <MAT2>

Query Match 44.8%; Score 40; DB 2; Length 2475;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 FPNKGTQQTDOE 16
DB 393 FLANKSTQANDPDE 406

RESULT 38

HAJSB

hemoglobin alpha-D chain - common starling

C:Species: Sturnus vulgaris (common starling)

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Mar-2000

C:Accession: A02331

R:Oberthur, W.; Braunitzer, G.

P:Seidler's Z. Physiol. Chem. 365, 159-173, 1984

Title: Hemoglobins of starling (Sturnus vulgaris, Passeriformes). The primary structure

Reference number: A91731; MID:84184202; PMID:6714943

A:Accession: A02331

A:Molecule type: protein

A:Residues: 1-141 <OBE>

C:Superfamily: globin; globin homology

C:Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier

F:2-141/Domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 43.8%; Score 39; DB 1; Length 141;

Best Local Similarity 53.3%; Pred. No. 32;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQI 15
DB 41 TYFPHPDLGQSDQI 55

RESULT 39

C24625

hemoglobin alpha-D chain - Eurasian tree sparrow

C:Species: Passer montanus (Eurasian tree sparrow)

C:Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 04-Mar-2000

C:Accession: C24625

R:Schneegans, D.; Braunitzer, G.; Oberthur, W.; Kostera, J.; Grimm, F.

P:Ol. Chem. Hoppe-Seyler 366, 893-899, 1985

Title: The hemoglobin of the tree sparrow (Passer montanus). The amino acid sequence

Reference number: A90696; MID:8607293; PMID:4074502

A:Accession: C24625

A:Molecule type: protein

A:Residues: 1-141 <SCH>

C:Superfamily: globin; globin homology

C:Keywords: chromoprotein; heme; iron; metalloprotein; oxygen carrier

F:2-141/Domain: globin homology <GLB>

F:58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 43.8%; Score 39; DB 2; Length 141;

Best Local Similarity 53.3%; Pred. No. 32;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQI 15
DB 41 TYFPHPDLGQSDQI 55

RESULT 40

T10407

Immediate early protein 0 - Orgyia pseudotsugata nuclear polyhedrosis virus

C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpMV

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000

C:Accession: T10407
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohmann, G.F.

Virology 229, 381-399, 1997

A:Title: The sequence of the Orgyia pseudotsugata multinnucleocapsid nuclear polyhedrosis

A:Reference number: Z17011; MID:97271300; PMID:9126251

A:Accession: T10407

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-245 <AHR>

A:Cross-references: EMBL:U75930; NID:G2934903; PID:G1911384

Query Match

Best Local Similarity 43.8%; Score 39; DB 2; Length 245;

Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQI 16
DB 5 THMPNIVSKSYTDACE 20

Search completed: August 20, 2003, 09:32:35
Job time : 18.9794 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 9.23711 Seconds

(without alignments)
81.457 Million cell updates/sec

Title: US-09-991-433-3

Perfect score: 89

Sequence: 1 TYFPNKGTOQYTDQIE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

al number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	781	1	COAT_PAVHB
2	43	48.3	547	1	CH60_PASMU
3	43	48.3	548	1	CH60_HAEIN
4	42	47.2	546	1	CH60_ACTPL
5	41	46.1	286	1	SPEE_BUCAT
6	40.5	45.5	349	1	DDI_LACTA
7	40	44.9	508	1	VGJG_HINV
8	40	44.9	725	1	GUNG_CLOCE
9	39	43.8	141	1	HEA2_CATMA
10	39	43.8	141	1	HBAD_PASMO
11	39	43.8	141	1	HBAD_STTVU
12	39	43.8	245	1	IBO_NPVOP
13	39	43.8	302	1	CANI_RABIT
14	39	43.8	306	1	PYRD_METJA
15	39	43.8	326	1	VS09_ROT99
16	39	43.8	398	1	SPEE_STRPY
17	39	43.8	458	1	NUAM_STRCA
18	39	43.8	546	1	VLI_HPV19
19	39	43.8	713	1	CANI_MOUSE
20	39	43.8	713	1	CANI_RAT
21	39	43.8	866	1	LOX1_LENCU
22	38	42.7	110	1	NUOK_RICCN
23	38	42.7	223	1	NUOB_BUCAP
24	38	42.7	324	1	Y365_BUCAT
25	38	42.7	414	1	IDHC_HUMAN
26	38	42.7	546	1	CH60_ACTAC
27	38	42.7	561	1	HCVB_PONLE
28	38	42.7	581	1	YG25_ARCFU
29	38	42.7	694	1	EPG_CHLMU
30	38	42.7	694	1	EPG_CHLTR
31	38	42.7	714	1	CANI_PIG
32	38	42.7	720	1	GSYL_BREDI
33	38	42.7	852	1	GSYL_YEAST

34	38	42.7	1066	1	HYGA_STRPN	Q54873 streptococ
35	38	42.7	2200	1	LAR_CAEEL	Q9bun8 caenorhabd
36	38	42.7	3396	1	PGCV_HUMAN	P13611 homo sapien
37	37	41.6	110	1	NUOK_RICPR	Q92922 rickettsia
38	37	41.6	161	1	YAS1_ARCFU	Q29211 archaeoglob
39	37	41.6	170	1	RS9_STRCO	Q53875 streptomyce
40	37	41.6	173	1	PACA_ONCNE	P41585 oncorhynchu
41	37	41.6	174	1	Y204_AQUAE	Q64400 aquifex aeo
42	37	41.6	228	1	B10D_VIBPA	Q87623 vibrio para
43	37	41.6	238	1	TRV3_SALSA	P35033 salmo salar
44	37	41.6	312	1	Y353_BUCAP	Q8K912 buchnera ap
45	37	41.6	414	1	IDHC_MOUSE	Q88844 mus musculu

ALIGNMENTS

```

RESULT 1
COAT_PAVHB          STANDARD;          PRT;          781 AA.
ID P07259;
AC 01-APR-1988 (rel. 07, Created)
DT 01-APR-1988 (rel. 07, Last sequence update)
DT 01-APR-1990 (rel. 14, Last annotation update)
DE Probable coat protein VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_Taxid=10798;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=Isolate AU;
RA MEDLINE=86200451; PubMed=3701931;
RA Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;
RT "Nucleotide sequence and genome organization of human parvovirus B19
   isolated from the serum of a child during aplastic crisis.";
RL J. Virol. 58:921-936(1986).
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, M13178; AAA66867.1; -.
CC PIR, A24299; VCPV19.
CC InterPro; IPR001403; Parvo coat.
CC DR Pfam; PR00740; Parvo coat; 1.
CC Coat protein; Glycoprotein.
CC FT CARBOHYD 46
CC FT CARBOHYD 46
CC FT CARBOHYD 184
CC FT CARBOHYD 220
CC FT CARBOHYD 293
CC FT CARBOHYD 293
CC SEQUENCE 781 AA; 86015 MW; 8C6254BD0576B07 CRC64;
CC -----
Query Match          100.0%;          Score 89;          DB 1;          Length 781;
Best Local Similarity 100.0%;          Pred. No. 1.5e-07;
Matches 16;          Conservative 0;          Mismatches 0;          Indels 0;          Gaps 0;
OY 1 TYFPNKGTOQYTDQIE 16
Db 618 TYFPNKGTOQYTDQIE 633
RESULT 2
CH60_PASMU          STANDARD;          PRT;          547 AA.
ID Q59687;
DT 01-NOV-1997 (rel. 35, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)

```

DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR GROEL OR MOPA OR PM1107.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 RX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PI059;
 RX MEDLINE=96105224; PubMed=8529887;
 RA Love B.C., Hansen L.M., Hirsch D.C.;
 RT "Cloning and sequence of the groESL heat-shock operon of Pasteurella
 multocida."
 RL Gene 166:179-180 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 proper assembly of unfolded polypeptides generated under stress
 conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 7 subunits (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U30165; AAA84916.1; -
 DR EMBL: AE006151; AAK03191.1; -
 DR PIR: JC4519; JC4519.
 DR HSSP: P06139; IJON.
 DR HAMAP: MF_00600; -; 1
 DR InterPro: IPR001844; Chaperin Cpn60.
 DR InterPro: IPR002423; Cpn60/TCF-1.
 DR Pfam: PF00118; cpn60_TCF1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding; Complete proteome.
 FT CONFLICT 424 424 S -> N (IN REF. 1).
 FT CONFLICT 464 464 I -> V (IN REF. 1).
 SQ SEQUENCE 547 AA; 57291 MW; EBB95BP235855E CRC64;
 Query Match 48.3%; Score 43; DB 1; Length 547;
 Best Local Similarity 66.7%; Pred. No. 9.3;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 5 NKGTQOYTDQIE 16
 Db 479 NAGTEQYGDWIE 490
 RESULT 3
 CH60_HABIN STANDARD; PRT; 548 AA.
 AC P43733;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR MOPA OR GROEL OR H10543.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.
 RX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Usterback T.R., Hanna W.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Gnehm C.L., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd".
 RL Science 269:496-512 (1995).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 proper assembly of unfolded polypeptides generated under stress
 conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 7 subunits (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32736; AAC2201.1; -
 DR PIR: C64076; C64076.
 DR HSSP: P06139; IJON.
 DR TIGR: H10543; -
 DR HAMAP: MF_00600; -; 1
 DR InterPro: IPR001844; Chaperin Cpn60.
 DR InterPro: IPR002423; Cpn60/TCF-1.
 DR Pfam: PF00118; cpn60_TCF1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 DR Chaperone; ATP-binding; Complete proteome.
 SQ SEQUENCE 548 AA; 57577 MW; CA4066AAC1B62159 CRC64;
 Query Match 48.3%; Score 43; DB 1; Length 548;
 Best Local Similarity 66.7%; Pred. No. 9.3;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 5 NKGTQOYTDQIE 16
 Db 480 NAGTEQYGDWIE 491
 RESULT 4
 CH60_ACTPL STANDARD; PRT; 546 AA.
 AC P94166;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR GROEL OR MOPA.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 RX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S 4074 / Serotype 1;

RA MEDLINE=97189570; PubMed=9037757;
RA Vezina G., Stiois M., Clairoux N., Boissinot M.;
RT "Cloning and characterization of the groE locus from *Actinobacillus*
RT *pleuropneumoniae*.";
RT FEMS Microbiol. Lett. 147:11-16(1997).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC-----
CC EMBL: U55016; AAB51437.1; -
CC
CC HSSP: P06139; IGRU.
CC
CC HAMAP: MF_00600; -; 1.
CC InterPro: IPR001844; Chaperlin Cpn60.
CC InterPro: IPR002423; Cpn60/TCP-1.
CC Pfam: PF00118; Cpn60 TCP1; 1.
CC PRINTS: PR00298; CHAPERONIN60.
CC PRINTS: PR00304; TCP1EXPTCP1.
CC PROSITE: PS00296; CHAPERONINS_CPN60; 1.
CC Chaperone; ATP-binding.
CC INIT MET 0
CC FT
CC SEQUENCE 546 AA; 57513 MW; E84BB72C9BD3DB56 CRC64;
CC
CC Query Match 47.2%; Score 42; DB 1; Length 546;
CC Best Local Similarity 53.3%; Pred. No. 14;
CC Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
CC
CC 2 YFPNKGTOGYTDQIE 16
CC Db 476 YGNAGTEQYGDMLE 490
CC
CC
CC RESULT 5
CC ID SPEE_BUCAI STANDARD; PRT; 286 AA.
CC AC P57305;
CC 16-OCT-2001 (Rel. 40; Created)
CC 16-OCT-2001 (Rel. 40; Last sequence update)
CC 28-FEB-2003 (Rel. 41; Last annotation update)
CC Spermidine synthase (EC 2.5.1.16) (Putrescine aminopropyltransferase)
CC (SPSY).
CC DE
CC SPEE OR B0209.
CC OS *Buchnera aphidicola* (subsp. *Acyrthosiphon pisum*) (*Acyrthosiphon pisum*
CC symbiotic bacterium).
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; *Buchnera*.
CC CX NCBI_Taxid=118099;
CC RN (1)
CC RP SEQUENCE FROM N.A.
CC RP STRAIN=Tokyo 1998;
CC RX MEDLINE=20445173; PubMed=10993077;
CC RX Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT *Buchnera* sp. APS.";
RT Buchnera 407:81-86(2000).
CC RL
CC Nature 407:81-86(2000).
CC -1- CATALYTIC ACTIVITY: S-adenosylmethioninamine + putrescine = 5'-
CC methylthioadenosine + spermidine.
CC -1- PATHWAY: Spermidine biosynthesis; fifth (last) step.
CC -1- SIMILARITY: BELONGS TO THE SPERMIDINE/SPERMINE SYNTHASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC-----

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AP001118; BAB12926.1; -
CC
CC HAMAP: MF_00198; -; 1.
CC InterPro: IPR000051; SAM bind.
CC InterPro: IPR001045; Sperm synthase.
CC Pfam: PF01564; Sperm_synth; 1.
CC TIGRFAMs: TIGR00417; speE; 1.
CC PROSITE: PS01330; SPERMIDINE SYNTHASE; 1.
CC Spermidine biosynthesis; Transferase; Complete proteome.
CC FT
CC DOMAIN 79 120
CC (POTENTIAL). BINDING TO DECARBOXYLATED SAM
CC
CC SEQUENCE 286 AA; 33788 MW; 5EE47D2107EB689E CRC64;
CC
CC Query Match 46.1%; Score 41; DB 1; Length 286;
CC Best Local Similarity 58.3%; Pred. No. 10;
CC Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
CC
CC 2 YFPNKGTOGYTD 13
CC Db 120 YFPNHSNOAYOD 131
CC
CC
CC RESULT 6
CC ID DDL_LACLA STANDARD; PRT; 349 AA.
CC AC Q9C1I5;
CC DT 16-OCT-2001 (Rel. 40; Created)
CC DT 16-OCT-2001 (Rel. 40; Last sequence update)
CC DT 15-SEP-2003 (Rel. 42; Last annotation update)
CC DE D-alanine--D-alanine ligase (EC 6.3.2.4) (D-alanylalanine synthetase)
CC GN DDL OR IL0341.
CC OS *Lactococcus lactis* (subsp. *lactis*) (*Streptococcus lactis*).
CC OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; *Lactococcus*.
CC CX NCBI_Taxid=1350;
CC RN (1)
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=IL1403;
CC RX MEDLINE=21235186; PubMed=11337471;
CC RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
RT Weisenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
RT *lactis* ssp. *lactis* IL1403.";
CC Genome Res. 11:731-753(2001).
CC -1- CATALYTIC ACTIVITY: CELL WALL FORMATION (BY SIMILARITY).
CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
CC alanyl-D-alanine.
CC -1- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
CC step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE006271; AAK04439.1; -
CC
CC PIR: E86667; E86667.
CC HSSP: P07662; 110M.
CC HAMAP: MF_00047; -; 1.
CC InterPro: IPR005905; D_ala_D_ala.
CC InterPro: IPR000291; Dala_lig_Van.
CC Pfam: PF01820; Dala_ligase; 1.
CC TIGRFAMs: TIGR01205; D_ala_lattGR; 1.
CC

Query Match 44.9%; Score 40; DB 1; Length 725;
Best Local Similarity 53.3%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Db 1 TYFPKNGTQYTDQI 15
271 STVPMKQGDQTDII 285

RESULT 9
ID_HBA2_CATMA STANDARD; PRT; 141 AA.
AC P82112;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hemoglobin alpha-2 chain (HB 2).
OS Catharacta maccomicki (South polar skua).
OC Archosauria; Aves; Neognathae; Charadriiformes; Stercorariidae;
Catharacta.
NCBI_TaxID=54055;
RN [1]
RP SEQUENCE.
RC TISSUE=Blood;
RX MEDLINE=20453132; PubMed=10998071;
RA Tamburini M., Riccio A., Romano M., Giardina B., di Prisco G.;
RT "Structural and functional analysis of the two hemoglobins of the
RT antarctic seabird Catharacta maccomicki. Characterization of an
RT additional phosphate binding site by molecular modelling.";
RL Eur. J. Biochem. 267:6089-6098(2000).
CC -1- FUNCTION: Involved in oxygen transport from the lung to the
CC various peripheral tissues.
CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
CC -1- TISSUE SPECIFICITY: Red blood cells.
CC -1- SIMILARITY: Belongs to the globin family.
CC HSSP; P02001; 1HBR.
DR InterPro; IPR002338; Alpha haem.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PR00612; ALPHAHAE.EM.
DR PROSITE; PS01033; GLOBIN; 1.
DR Heme; Oxygen transport; Transport; Erythrocyte.
KW METAL 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15930 MW; F535256589083C4 CRC64;
Query Match 43.8%; Score 39; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Db 1 TYFPKNGTQYTDQI 15
41 TYFPKNGTQYTDQI 55

RESULT 10
ID_HBA2_CATMA STANDARD; PRT; 141 AA.
AC P82112;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Hemoglobin alpha-D chain.
OS Passer montanus (Tree sparrow).
OC Archosauria; Aves; Neognathae; Charadriiformes; Stercorariidae;
Catharacta.
NCBI_TaxID=9160;
RN [1]
RP SEQUENCE.
RC TISSUE=Blood;
RX MEDLINE=20453132; PubMed=10998071;
RA Tamburini M., Riccio A., Romano M., Giardina B., di Prisco G.;
RT "Structural and functional analysis of the two hemoglobins of the
RT antarctic seabird Catharacta maccomicki. Characterization of an
RT additional phosphate binding site by molecular modelling.";
RL Eur. J. Biochem. 267:6089-6098(2000).
CC -1- FUNCTION: Involved in oxygen transport from the lung to the
CC various peripheral tissues.
CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
CC -1- TISSUE SPECIFICITY: Red blood cells.
CC -1- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
CC hemoglobin component, called hemoglobin d, which is expressed in
CC late embryonic and adult life.
CC -1- SIMILARITY: Belongs to the globin family.
CC PIR; A02331; HAI5B.
DR HSSP; P02001; 1HBR.
DR InterPro; IPR002338; Alpha haem.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PR00612; ALPHAHAE.EM.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT METAL 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15660 MW; 647C64086BC97D89 CRC64;
Query Match 43.8%; Score 39; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Db 1 TYFPKNGTQYTDQI 15
41 TYFPKNGTQYTDQI 55

Query Match 43.8%; Score 39; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 10;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTQOYTDQIE 15
 |||||
 DB 41 THWPNLVSKSYTDACE 20

RESULT 12
 IEQ_NPVOP STANDARD; PRT; 245 AA.
 AC 010369;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Immediate-early protein IE-0.
 GN IE-0.
 OS Oryzias pseudoscutigata multicapsid polyhedrosis virus (OpMVV).
 VRuses; dsDNA viruses, no RNA stage; Baculoviridae;
 Nucleopolyhedrovirus.
 NCBI_TaxId=164623;
 RN [1]
 RX MEDLINE=97271300; PubMed=9126251;
 RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
 Rohmann G.F.;
 RT "The sequence of the Oryzias pseudoscutigata multinucleocapsid nuclear
 polyhedrosis virus genome.";
 RL Virology 229:381-399(1997).
 CC -1 SIMILARITY: Contains 1 RING-type zinc finger.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U75930; AAC59137.1; -.
 DR InterPro: IPR001841; Zn_finger.
 DR Pfam: PF005290; Baculo_IE-1; 1.
 DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
 DR PROSITE: PS00089; ZF_RING_2; 1.
 DR KW Early protein; zinc-finger.
 FT ZN RING 195 240 RING-TYPE.
 FT SEQUENCE 245 AA; 27117 MW; 357F43B15F7B1029 CRC64;
 QY Query Match 43.8%; Score 39; DB 1; Length 245;
 Best Local Similarity 43.8%; Pred. No. 19;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 TYFPNKGTQOYTDQIE 16
 |||||
 DB 5 THWPNLVSKSYTDACE 20

RESULT 13
 CANI_RABIT STANDARD; PRT; 302 AA.
 ID CANI_RABIT
 AC P06815;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated
 DE neutral protease) (CANP) (Mu-type) (mucAMP) (micromolar-calpain)
 DE (Frogmont).
 DE CAPN1.
 GN Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxId=9986;

RN [1]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=86250902; PubMed=2424911;
 RA Emori Y., Kawasaki H., Sugihara H., Imaish S., Kawashima S.,
 RA Suzuki K.;
 RT "Isolation and sequence analyses of cDNA clones for the large
 RT subunit of two isozymes of rabbit calcium-dependent protease.";
 RL J. Biol. Chem. 261:9465-9471(1986).
 RN [2]
 RP CALCIUM-BINDING DATA.
 RX MEDLINE=87279982; PubMed=3038855;
 RA Minami Y., Emori Y., Kawasaki H., Suzuki K.;
 RT "B-F hand structure-domain of calcium-activated neutral protease
 RT (CANP) can bind Ca²⁺ ions.";
 RL J. Biochem. 101:889-895(1987).
 CC -1 FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction.
 CC -1 CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the p2 residue.
 CC -1 COFACTOR: Binds 3 calcium ions (By similarity).
 CC -1 ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1 SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca²⁺ binding (By similarity).
 CC -1 SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M1363; AAA11456.1; -.
 DR PIR: A24815; A24815.
 DR HSP: Q07009; 1DF0.
 DR MEROPS: C02.001; -.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001300; Protease C2.
 DR InterPro: IPR000169; SH3prot_acetate.
 DR Pfam: PF01067; Calpain_III; 1.
 DR Pfam: PF00036; ehfand_3; 1.
 DR SMART: SM00720; calpain_III; 1.
 DR SMART: SM00054; Eph; 3.
 DR PROSITE: PS00018; EF_HAND; 2.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; PARTIAL.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; PARTIAL.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; PARTIAL.
 DR KW Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
 FT NON_TER 1 114
 FT DOMAIN 1 114
 FT DOMAIN 115 130
 FT DOMAIN 131 301
 FT CA_BIND 142 153
 FT CA_BIND 186 197
 FT CA_BIND 216 227
 FT DOMAIN 251 262
 FT DOMAIN 281 292
 FT SEQUENCE 302 AA; 35274 MW; 178BPF4216CEAB CRC64;
 QY Query Match 43.8%; Score 39; DB 1; Length 302;
 Best Local Similarity 33.3%; Pred. No. 24;
 Matches 10; Conservative 3; Mismatches 3; Indels 14; Gaps 1;

QY 1 TYFPNK-----GTQOYTDQIE 16
 |||||
 DB 88 TFPNKGDPVLRFPSEKRAQTQIEDQIQ 117

RESULT 14
PYRD METUA STANDARD; PRT; 306 AA.
ID AC 058070;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydroorotate dehydrogenase (EC 1.3.3.1) (Dihydroorotate oxidase)
GN PYRD OR M0654.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=868087;
But C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RT Science 273:1058-1073(1996).
RL -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) = orotate +
H(2)O(2).
CC -1- COFACTOR: FMN (BY SIMILARITY).
CC -1- PATHWAY: Pyrimidine biosynthesis; fourth step.
CC -1- SUBUNIT: Heterotrimer of 2 pyrK and 2 pyrd subunits (By
similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the dihydroorotate dehydrogenase family.
Subfamily 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U67513; AAB98649.1; -
PIR; F64381; F64381.
HSSP; P54321; 2DOR.
DR TIGR; M0654; -
DR HAMAP; MF_00224; -; 1.
DR InterPro; IPR001295; DHO_dh.
DR InterPro; IPR005720; DHO_dhl.
DR InterPro; IPR003009; FMO enzyme.
DR Pfam; PF01180; DHOdehase; 1.
DR TIGRFAMs; TIGR01037; pyrd sub1_fam; 1.
DR PROSITE; PS00911; DHOdehase_1; 1.
DR PROSITE; PS00912; DHOdehase_2; 1.
KW Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FMN;
KW Complete proteome.
FT NP BIND 240 FMN (POTENTIAL).
SQ SEQUENCE 306 AA; 32940 MW; E5352389BA9FC7C3 CRC64;

Query Match 43.8%; Score 39; DB 1; Length 306;
Best Local Similarity 46.2%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DB 76 PNPQVDTLLEIE 88

RESULT 15
VS09 ROT99 STANDARD; PRT; 326 AA.
ID AC P33492;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycoprotein VP7 (Serotype-specific antigen) (Outer shell
glycoprotein).
GN S9.
OS Human rotavirus (serotype G / strain RK9).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=33722;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91214432; PubMed=1965285;
RX Besarab I.N., Novikova P.A., Borodin A.M.;
RT "Use of the polymerase chain reaction for analysis of rotaviruses.
Nucleotide sequence of a gene, coding for the basic neutralizing
antigen VP7 of a human rotavirus with a new G-serotype (letter).";
RL Bioorg. Khim. 16:1689-1691(1990).
CC -1- FUNCTION: VP7 IS THE MAJOR NEUTRALIZATION ANTIGEN.
CC -1- SUBCELLULAR LOCATION: Rough endoplasmic reticulum.
DR PIR; J00307; J00307.
DR InterPro; IPR001963; VP7.
DR Pfam; PF00434; VP7; 1.
DR ProDom; PD000191; VP7; 1.
DR Coat protein; Transmembrane; Glycoprotein.
FT TRANSMEM 32 48
FT CARBOHYD 69 69 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 37323 MW; 3C243914E23A8FE CRC64;

Query Match 43.8%; Score 39; DB 1; Length 326;
Best Local Similarity 58.3%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DB 84 YFNDVNTQITD 95

RESULT 16
SEB STREPY STANDARD; PRT; 398 AA.
ID AC P00788; P26296; Q54960; Q54961; Q54963; Q54964; Q54965;
AC Q54966; Q54967; Q54968; Q57024; Q57082; Q57202; Q57211; Q57212;
AC Q9S680;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Streptopain precursor (EC 3.4.22.10) (Streptococcal cysteine
proteinase) (Streptococcus peptidase A) (SPe) (Exotoxin type B)
GN SPeB OR SPY2039 OR SPYM3_1742 OR SPS1739 OR SPYM18_2099.
OS Streptococcus pyogenes.
OS Streptococcus pyogenes (serotype M3) and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466, 186103;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-32 AND 146-162.
RC STRAIN=86-858, and NY-5;
RX MEDLINE=90330563; PubMed=2198264;
RA Hausser A.R., Schlevert P.M.;
RT "Nucleotide sequence of the streptococcal pyrogenic exotoxin type B
gene and relationship between the toxin and the streptococcal
proteinase precursor.";
RT U. Bacteriol. 172:4536-4542(1990).
RL [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=789 / Serotype M1, 327 / Serotype M2, 75 / Serotype M3,

RC 157 / Serotype M3, 315 / Serotype M3, 1251 / Serotype M3,
 RC 321 / Serotype M4, 1289 / Serotype M5, 303 / Serotype M6,
 RC 429 / Serotype M8, 796 / Serotype M9, 800 / Serotype M9,
 RC 1896 / Serotype M10, 650 / Serotype M11, 282 / Serotype M12,
 RC 659 / Serotype M13, 660 / Serotype M14, 198 / Serotype M15,
 RC 1233 / Serotype M17, 156 / Serotype M18, 300 / Serotype M18,
 RC 1294 / Serotype M19, 162 / Serotype M22, 1901 / Serotype M23,
 RC 684 / Serotype M24, 686 / Serotype M25, 1838 / Serotype M27,
 RC 587 / Serotype M28, 366 / Serotype M30, 427 / Serotype M31,
 RC 807 / Serotype M33, 1841 / Serotype M41, 1842 / Serotype M43,
 RC 1226 / Serotype M44, 719 / Serotype M49, 1864 / Serotype M56,
 RC 1882 / Serotype M59, 168 / Serotype M66, 302 / Serotype M73,
 RC 758 / Serotype M75, 1911 / Serotype M76, 1832 / Serotype M76,
 RC 165 / Serotype M, 262 / Serotype M, 317 / Serotype M,
 RC 1990 / Serotype M, 1991 / Serotype M, 2017 / Serotype M,
 RC 2018 / Serotype M, 1719 / Serotype M, 289 / Serotype M,
 RC 1590, 1870, 1871, 1872, 1893, and 1914A;
 MEDLINE=94285789; PubMed=7516997;
 Kapur V., Topouzis S., Majeeky M.W., Li L.L., Hamrick M.R.,
 Hamill R.J., Patti J.M., Musser J.M.,
 "A conserved Streptococcus pyogenes extracellular cysteine protease
 cleaves human fibronectin and degrades vitronectin.",
 Microb. Pathog. 15:327-346(1993).
 [3]
 RA SEQUENCE FROM N.A.
 RC STRAIN=A-20 / Serotype M1,T1;
 RA Wu J.-J.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 [4]
 RA SEQUENCE FROM N.A.
 RC STRAIN=SV / Serotype M23;
 RA Hong K.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 [5]
 RA SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 MEDLINE=21192684; PubMed=11296296;
 Ferrerelli J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 Primeaux C., Szatze S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 Qian X., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 "Complete genome sequence of an M1 strain of Streptococcus
 pyogenes.",
 Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 [6]
 RA SEQUENCE FROM N.A.
 RC STRAIN=MGAS315 / Serotype M3;
 MEDLINE=22133808; PubMed=12122206;
 Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 Schlievert P.M., Musser J.M.;
 "Genome sequence of a serotype M3 strain of group A Streptococcus:
 phase-encoded toxins, the high-virulence phenotype, and clone
 emergence.",
 Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 [7]
 RA SEQUENCE FROM N.A.
 RC STRAIN=SSI-1 / Serotype M3;
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 Yamazaki K., Okahashi N., Kawabata S., Yaunaga T., Hattori M.,
 Hayashi H., Hamada S.;
 "The genome of invasive Streptococcus pyogenes; a comparative analysis
 of S. pyogenes SSI-1, SF370 and MGAS8232.",
 Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 [8]
 RA SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,

RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 "Genome sequence and comparative microarray analysis of serotype M18
 group A Streptococcus strains associated with acute rheumatic fever
 outbreaks.",
 Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 [9]
 RA PRELIMINARY SEQUENCE OF 28-86 AND 121-398.
 RA Yonaha K., Elliott S.D., Liu T.-Y.;
 "Primary structure of zymogen of streptococcal proteinase.",
 J. Protein Chem. 1:317-334(1982).
 [10]
 RA PRELIMINARY SEQUENCE OF 146-398.
 RX MEDLINE=76190087; PubMed=1270417;
 Tai J.Y., Korte A.A., Liu T.-Y., Elliott S.D.;
 "Primary structure of streptococcal proteinase. III. Isolation of
 cyanogen bromide peptides: complete covalent structure of the
 polypeptide chain.",
 J. Biol. Chem. 251:1955-1959(1976).
 [11]
 RA FUNCTION.
 RP STRAIN=NZ131 / Serotype M49,T14;
 RX MEDLINE=99081733; PubMed=9864206;
 Kuo C.-F., Wu J.-J., Tsai P.-J., Kao F.-J., Lei H.-Y., Lin M.T.,
 Lin Y.-S.;
 "Streptococcal pyrogenic exotoxin B induces apoptosis and reduces
 phagocytic activity in U937 cells.",
 Infect. Immun. 67:126-130(1999).
 [12]
 RA FUNCTION.
 RP STRAIN=A-20 / Serotype M1,T1;
 RX MEDLINE=99386817; PubMed=10456871;
 Tsai P.-J., Lin Y.-S., Kuo C.-F., Lei H.-Y., Wu J.-J.;
 "Group A Streptococcus induces apoptosis in human epithelial cells.",
 Infect. Immun. 67:4334-4339(1999).
 [13]
 RA FUNCTION: IMPORTANT STREPTOCOCCAL VIRULENCE FACTOR WHICH CLEAVES
 HUMAN FIBRONECTIN AND DEGRADES VITRONECTIN. ALSO CLEAVES HUMAN
 IL1B PROCURSOR TO FORM BIOLOGICALLY ACTIVE IL1B. CAN INDUCE
 APOPTOSIS IN HUMAN MONOCYTES AND EPITHELIAL CELLS IN VITRO, AND
 REDUCES PHAGOCYTIC ACTIVITY IN MONOCYTIC CELLS. THUS, MAY PLAY A
 ROLE IN BACTERIAL COLONIZATION, INVASION, AND INHIBITION OF WOUND
 HEALING.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage with hydrophobic
 CC residues at p2, p1 and p1'.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE
 CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET
 CC FEVER, HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE
 CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC
 CC FEVER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL; M86905; AAA26978.1; -
 DR EMBL; L26125; AAA26979.1; -
 DR EMBL; L26126; AAA26982.1; -
 DR EMBL; L26127; AAA26993.1; -
 DR EMBL; L26128; AAA26994.1; -
 DR EMBL; L26130; AAA26996.1; -
 DR EMBL; L26131; AAA26997.1; -
 DR EMBL; L26132; AAA26998.1; -
 DR EMBL; L26133; AAA26999.1; -
 DR EMBL; L26134; AAA27000.1; -
 DR EMBL; L26135; AAA27001.1; -
 DR EMBL; L26136; AAA27002.1; -
 DR EMBL; L26137; AAA27003.1; -
 DR EMBL; L26138; AAA27004.1; -

DR EMBL, L26139; AAA27005.1; -
DR EMBL, L26140; AAA27006.1; -
DR EMBL, L26141; AAA27007.1; -
DR EMBL, L26142; AAA27008.1; -
DR EMBL, L26143; AAA27009.1; -
DR EMBL, L26144; AAA27010.1; -
DR EMBL, L26145; AAA27011.1; -
DR EMBL, L26146; AAA27012.1; -
DR EMBL, L26147; AAA27013.1; -
DR EMBL, L26148; AAA27014.1; -
DR EMBL, L26149; AAA27015.1; -
DR EMBL, L26150; AAA27016.1; -
DR EMBL, L26151; AAA26980.1; -
DR EMBL, L26152; AAA26981.1; -
DR EMBL, L26153; AAA26982.1; -
DR EMBL, L26154; AAA26983.1; -
DR EMBL, L26155; AAA26984.1; -
DR EMBL, L26156; AAA26985.1; -
DR EMBL, L26157; AAA26986.1; -
DR EMBL, L26158; AAA26987.1; -
DR EMBL, L26159; AAA26988.1; -
DR EMBL, L26160; AAA26989.1; -
DR EMBL, L26161; AAA26990.1; -
DR EMBL, L26162; AAA26991.1; -
DR EMBL, AF104940; AAD17930.1; -

Query Match 43.8% Score 39; DB 1; Length 398;
Best Local Similarity 60.0%; Pred. No. 32;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 PENKGTQOYT 12
: ||| : |||
: ||| : |||

Db 206 YENKGLKDYT 215

RESULT 17
ID NU4M_STRCA STANDARD; PRT; 458 AA.

AC 021406;

DT 15-JUL-1998 (Rel. 36, Created)

DT 28-EB-2003 (Rel. 41, Last annotation update)

DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).

GN MTND4 OR ND4 OR NADH4.

OS Struthio camelus (Ostrich).

OC Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;

OC Struthio.

NCBI_TaxID=8801;

RM [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=97357422; PubMed=9214748;

RA Harlid A., Janke A., Arnason U.;

RT "The mtDNA sequence of the ostrich and the divergence between

RT paleognathous and neognathous birds.";

RL Mol. Biol. Evol. 14:754-761(1997).

RM [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=21263106; PubMed=11370967;

RA Haddrath O., Baker A.J.;

RT "Complete mitochondrial DNA genome sequences of extinct birds: rattle

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC EMBL, Y12025; CAK27253.1; -
DR EMBL, AF338715; AAK53354.1; -
DR PIR, B99613; B99613.
DR PIR, T11528; T11528.
DR InterPro; IPR003918; NADHub_oxred4.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR000260; Oxidored_q5_N.
DR Pfam; PF00361; oxidored_q5_N.1.
DR Pfam; PF01059; oxidored_q5_N.1.
DR PRINTS; PR01437; NUOXDRDTASE4.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
KW SEQUENCE 458 AA; 50955 MW; 575118B6297A18AC CRC64;

Query Match 43.8% Score 39; DB 1; Length 458;
Best Local Similarity 52.9%; Pred. No. 38;

Matches 9; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

Qy 1 TYFPNKGTQOYT--DOI 15
: ||| : |||
: ||| : |||

Db 45 TYFPKNNMTPTWGTGDIQI 61

RESULT 18
ID VLI_HPV19 STANDARD; PRT; 546 AA.

AC 002050;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Major capsid protein L1.

GN L1.

OS Human papillomavirus type 19.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

NCBI_TaxID=10608;

RM [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=94265501; PubMed=8205838;

RA Delius H., Hofmann B.;

RT "Primer-directed sequencing of human papillomavirus types.";

RT Curr. Top. Microbiol. Immunol. 186:13-31(1994).

RL J. Virol. 66:5714-5725(1992).

RM [2]

RP SEQUENCE OF 344-387 FROM N.A.

RA MEDLINE=92407963; PubMed=1326639;

RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;

RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and

RT variants: a showcase for the molecular evolution of DNA viruses.";

RL J. Virol. 66:5714-5725(1992).

RM [2]

RP SEQUENCE OF 344-387 FROM N.A.

RA MEDLINE=92407963; PubMed=1326639;

RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;

RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and

RT variants: a showcase for the molecular evolution of DNA viruses.";

RL J. Virol. 66:5714-5725(1992).

RM [2]

RP SEQUENCE OF 344-387 FROM N.A.

RA MEDLINE=92407963; PubMed=1326639;

RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;

RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and

RT variants: a showcase for the molecular evolution of DNA viruses.";

RL J. Virol. 66:5714-5725(1992).

RM [2]

RP SEQUENCE OF 344-387 FROM N.A.

Db 313 TYXIPNNSSQOQYTN 327

RESULT 19
CAN1_MOUSE STANDARD; PRT; 713 AA.

AC O35350; O88666;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated neutral proteinase) (CAMP) (Mu-type) (mucAMP) (micromolar-calpain).
GN CANP1 OR CANP1 OR CANP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
SEQUENCE FROM N.A.

Poirier C., Poussard S., Faust D.M., Imaizumi-Shearer T., Weiss M.C., Ducastang A., Montarras D., Pinet C., Guenet J.-L.;
Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

STRAIN=129;

Sahr K.E., Andrabi S., Peters L.L., Chishti A.H.;

"Cloning and characterization of the cDNA and gene encoding the mouse mu-calpain large subunit protein."
Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which catalyzes limited proteolysis of substrates involved in cytoskeletal remodelling and signal transduction.

-1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or Arg|-Xaa with Leu or Val as the P2 residue.

-1- COFACTOR: Binds 3 calcium ions (By similarity).

-1- ENZYME REGULATION: Activated by micromolar concentrations of calcium and inhibited by calpastatin.

-1- SUBUNIT: Heterodimer of a large (catalytic) and a small (regulatory) subunit.

-1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma membrane upon Ca++ binding (By similarity).

-1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF021847; AAB72222.1; -;
EMBL; AF084459; AAC3134.1; -;
HSSP; O07009; IDPO.

MEROPS; C02.001; -;

MCD; MGT; 88263; Canp1.

InterPro; IPR002048; EF-hand.

InterPro; IPR001300; Protease_C2.

Pfam; PF01067; Calpain III; 1.

Pfam; PF00036; ehand; 3.

Pfam; PF00648; Peptidase_C2; 1.

PRINTS; PR00704; CALPAIN.

SMART; SM00720; calpain III; 1.

SMART; SM00054; EFP; 3.

PROSITE; PS00018; EF_HAND; 2.

PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.

PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.

PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.

KW Hydrolase; Thiol protease; Calcium-binding; Multigene family.

FT DOMAIN 30 365 CALPAIN.
FT DOMAIN 366 525 CALPAIN III.
FT DOMAIN 526 541 LINKER.
FT DOMAIN 542 712 DOMAIN IV.
FT CA_BIND 553 564 EF-HAND 1.
FT CA_BIND 597 608 EF-HAND 2.
FT CA_BIND 627 638 EF-HAND 3.
FT DOMAIN 662 733 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT ACT_SITE 692 703 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
FT ACT_SITE 115 115 BY SIMILARITY.
FT ACT_SITE 212 212 BY SIMILARITY.
FT ACT_SITE 296 296 BY SIMILARITY.
FT CONFLICT 488 488 L -> P (IN REF. 2).
FT CONFLICT 696 696 D -> N (IN REF. 2).
SQ SEQUENCE 713 AA; 82106 MW; 3E1E26C95802B864 CRC64;

Query Match 43.8%; Score 39; DB 1; Length 713;
Best Local Similarity 33.3%; Pred. No. 61;
Matches 10; Conservative 3; Mismatches 3; Indels 14; Gaps 1;

QY 1 TYFPNK-----GTQOYTDQIE 16
Db 499 TREPNNKEDFLRFPSEKKAQTQELDDQIQ 528

RESULT 20

CAN1_RAT STANDARD; PRT; 713 AA.

AC P97571;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated neutral proteinase) (CAMP) (Mu-type) (mucAMP) (micromolar-calpain).

GN CAPN1 OR CUS1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxId=10116;

RN (1)

RP SEQUENCE FROM N.A.

RA MEDLINE=97107433; PubMed=8950173;

RA Sorimachi H., Amano S., Ishiura S., Suzuki K.;

RT "Primary sequences of rat mu-calpain large and small subunits are, respectively, moderately and highly similar to those of human.";

RL Biochim. Biophys. Acta 1309:37-41(1996).

-1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which catalyzes limited proteolysis of substrates involved in cytoskeletal remodelling and signal transduction.

-1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or Arg|-Xaa with Leu or Val as the P2 residue.

-1- COFACTOR: Binds 3 calcium ions (By similarity).

-1- ENZYME REGULATION: Activated by micromolar concentrations of calcium and inhibited by calpastatin.

-1- SUBUNIT: Heterodimer of a large (catalytic) and a small (regulatory) subunit.

-1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma membrane upon Ca++ binding (By similarity).

-1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U53858; AAC53001.1; -;
PDB; 1KKR; 20-MAR-02.

MEROPS; C02.001; -;

DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR000169; SHProt acetate.
 DR Pfam; PF000367; Calpain_III; 1.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Calpsec; 1.
 DR SMART; SM00054; Eph; 3.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KW Calcium-binding; Hydroxylase; Thiol protease; 3d-structure.
 FT DOMAIN 30 365 CALPAIN
 FT DOMAIN 366 525 DOMAIN_III.
 FT DOMAIN 526 541 LINKER.
 FT DOMAIN 542 712 DOMAIN_IV.
 FT CA_BIND 553 564 EF_HAND_1.
 FT CA_BIND 597 608 EF_HAND_2.
 FT CA_BIND 627 638 EF_HAND_3.
 FT DOMAIN 662 673 ANCESTRAL_CALCITUM_SITE_3 (POTENTIAL).
 FT DOMAIN 692 703 ANCESTRAL_CALCITUM_SITE_4 (POTENTIAL).
 FT ACT_SITE 115 115 BY SIMILARITY.
 FT ACT_SITE 272 272 BY SIMILARITY.
 FT ACT_SITE 296 296 BY SIMILARITY.
 SQ SEQUENCE 713 AA; 82118 MW; 6E66460B0BFAPBB CRC64;

Query Match 43.8%; Score 39; DB 1; Length 713;
 Best Local Similarity 33.3%; Pred. No. 61;
 Matches 10; Conservative 3; Mismatches 3; Indels 14; Gaps 1;

QY 1 TYPKPK-----GTQQTDPQIE 16
 Db 499 TYPKPKGDFLLRFSEKKAQTGLDDQIQ 528

RESULT 21
 LOX1_LENCU STANDARD; PRT; 866 AA.
 AC P38414;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lipoxigenase (EC 1.13.11.12).
 GN LOX1.
 Lens culinaris (lentil).
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 eurosida I; Fabales; Fabaceae; Papilionoideae; Viciae; Lens.
 OC NCBI_TaxID=3864;
 RX MEDLINE=94162308; PubMed=8117753;
 RA Hilbers M.P., Roosi A., Finazzi-Agro A., Veldink G.A.,
 RA Vliegenhart J.F.G.;
 RT "The primary structure of a lipoxigenase from the shoots of etiolated
 lentil seedlings derived from its cDNA."
 RL Biochim. Biophys. Acta 1211:239-242 (1994).
 CC -!- FUNCTION: PLANT LIPOXYGENASE MAY BE INVOLVED IN A NUMBER OF
 DIVERSE ASPECTS OF PLANT PHYSIOLOGY INCLUDING GROWTH AND
 DEVELOPMENT, PEST RESISTANCE, AND SENSENCE OR RESPONSES TO
 WOUNDING. IT CATALYZES THE HYDROPEROXIDATION OF LIPIDS, CONTAINING
 A CIS, CIS-1,4-PENTADIENE STRUCTURE.
 CC -!- CATALYTIC ACTIVITY: lipoate + O(2) = (9Z,11E)-(13S)-13-
 hydroperoxyoctadeca-9,11-dienoate.
 CC -!- COFACTOR: IRON, ONE ATOM TIGHTLY BOUND PER MOLECULE.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE LIPOXYGENASE FAMILY.
 CC -!- SIMILARITY: Contains 1 PLAT domain.

CC -----PROT entry is copyright. It is produced through a collaboration
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; X71344; CA50483.1; -.
 DR HSSP; P08170; ZSBL.
 DR InterPro; IPR000907; Lipoxigenase.
 DR InterPro; IPR001024; Lipoxigenase_LH2.
 DR Pfam; PF00305; lipoxigenase; 1.
 DR Pfam; PF01477; PLAT; 1.
 DR PRINTS; PR00087; LIPOXYGENASE.
 DR SMART; SM00308; LH2; 1.
 DR PROSITE; PS00711; LIPOXYGENASE_1; 1.
 DR PROSITE; PS0081; LIPOXYGENASE_2; 1.
 DR PROSITE; PS00095; PLAT; 1.
 DR Oxidoreductase; Dioxygenase; Iron; Multigene family.
 FT DOMAIN 42 171 PLAT.
 FT METAL 526 526 IRON (BY SIMILARITY).
 FT METAL 531 531 IRON (BY SIMILARITY).
 FT METAL 717 717 IRON (BY SIMILARITY).
 FT METAL 866 866 IRON (BY SIMILARITY).
 SQ SEQUENCE 866 AA; 96638 MW; E3B5D0B84E3C89C CRC64;

Query Match 43.8%; Score 39; DB 1; Length 866;
 Best Local Similarity 50.0%; Pred. No. 76;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPKPKGTQGYTD 13
 Db 741 FMPKGTPEYTD 752

RESULT 22
 NIOX_RICCN STANDARD; PRT; 110 AA.
 AC Q92698;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADH-quinone oxidoreductase chain K (EC 1.6.99.5) (NADH dehydrogenase
 D1, chain K) (NDH-1, chain K).
 GN NUOK OR RC1225.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OC NCBI_TaxID=781;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii".
 RL Science 293:2093-2098 (2001).
 CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
 sulfur (Fe-S) centers, to quinones in the respiratory chain.
 CC Couples the redox reaction to proton translocation (for every two
 CC electrons transferred, four hydrogen ions are translocated across
 CC the cytoplasmic membrane), and thus conserves the redox energy in
 CC a proton gradient (By similarity).
 CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: TO POLYPEPTIDE 4L OF THE NADH-UBIQUINOL OXIDOREDUCTASE
 CC OF CHLOROPLASTS OR MITOCHONDRIA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; A008669; AAL03763.1; -

DR PIR; A97853; A97853.

DR InterPro; IPR003215; NADH_dh_ubiq1.

DR InterPro; IPR001133; Oxidored4L.

DR InterPro; IPR003214; Oxidored4L.

DR Pfam; PF00420; Oxidored_q2; 1.

DR ProDom; PD002107; NADH_dh_ubiq1; 1.

CC Oxidoreductase; NAD; Quinone; Transmembrane; Complete proteome.

FT TRANSMEM 13 33 POTENTIAL.

FT TRANSMEM 41 61 POTENTIAL.

FT TRANSMEM 73 93 POTENTIAL.

SQ SEQUENCE 110 AA; 12535 MW; A2B99F11E4B5F266 CRC64;

Query Match 42.7%; Score 38; DB 1; Length 110;
Best Local Similarity 58.3%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YFPNKGTOGYTD 13

DB 93 YFPNKGSIETD 104

RESULT 23

NUOB_BUCAP STANDARD; PRT; 223 AA.

ID NUOB_BUCAP

AC 08K9Y6;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE NADH-quinone oxidoreductase chain B (EC 1.6.99.5) (NADH dehydrogenase

I, chain B) (NDH-1, chain B).

GN NUOB OR BUSG148.

OS Buchnera aphidicola (subsp. Schizaphis graminum).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Buchnera.

OX NCBI_TaxID=98794;

(1)

SEQUENCE FROM N.A.

RA MEDLINE=2084549; PubMed=12089438;

Tamara I., Klasson L., Candaeck B., Naeelund A.K., Eriksson A.-S.,

Wernegren J.V., Sandstrom J.P., Moran N.A., Andersson S.G.B.;

"50 million years of genomic stasis in endosymbiotic bacteria.";

Science 296:2376-2379(2002).

-1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-

sulfur (Fe-S) centers, to quinones in the respiratory chain.

Couples the redox reaction to proton translocation (for every two

electrons transferred, four hydrogen ions are translocated across

the cytoplasmic membrane), and thus conserves the redox energy in

a proton gradient (By similarity).

-1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.

-1- COFACTOR: May contain an iron-sulfur 4Fe-4S cluster.

-1- SUBUNIT: Composed of 13 different subunits (By similarity).

-1- SIMILARITY: BELONGS TO THE COMPLEX I 20 KDa SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; A014090; AAM67716.1; -

DR InterPro; IPR006138; Cmpixl_20kda.

DR InterPro; IPR006137; Oxidored_96.

DR Pfam; PF01058; Oxidored_96; 1.

DR PROSITE; PS01150; COMPLEX1_20K; 1.

KW Oxidoreductase; NAD; Quinone; Iron-sulfur; 4Fe-4S;

KM Complete proteome.

FT METAL 66 66 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 67 67 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 132 132 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 161 161 IRON-SULFUR (4FE-4S) (POTENTIAL).

SQ SEQUENCE 223 AA; 25497 MW; 9883840771B302BF CRC64;

Query Match 42.7%; Score 38; DB 1; Length 223;

Best Local Similarity 46.7%; Pred. No. 26;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YFPNKGTOGYTDQIE 16

DB 15 HYFNKTOSTIDPIE 29

RESULT 24

Y365_BUCAI STANDARD; PRT; 324 AA.

ID Y365_BUCAI

AC P57446;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein BU365.

GN BU365.

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

synbiotic bacterium).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Buchnera.

OX NCBI_TaxID=118099;

(1)

SEQUENCE FROM N.A.

RP STRAIN=Tokyo 1998;

RX MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

"Genome sequence of the endocellular bacterial symbiont of aphids

Buchnera sp. APS.";

RL Nature 407:81-86(2000).

-1- SIMILARITY: BELONGS TO THE UPF0176 FAMILY.

-1- SIMILARITY: Contains 1 rhodanese domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AP001119; BAB13069.1; -

DR HAMAP; MF_00469; -; 1.

DR InterPro; IPR001763; Rhodanese-like.

DR Pfam; PF00581; Rhodanese; 1.

DR SMART; SM00450; RHOD; 1.

DR PROSITE; PS0206; RHODANSE 3; 1.

KW Hypothetical protein; Complete proteome.

FT DOMAIN 145 239 RHODANSE.

SQ SEQUENCE 324 AA; 38405 MW; 9E7B019857D377FA CRC64;

Query Match 42.7%; Score 38; DB 1; Length 324;

Best Local Similarity 50.0%; Pred. No. 39;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPNKGTOGYTDQI 15

DB 32 YFPINOTBYRDRL 45

RESULT 25

IDHC_HUMAN STANDARD; PRT; 414 AA.

ID IDHC_HUMAN

AC Q75874; Q93090; Q9NT99; Q9UKW8;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Isocitrate dehydrogenase [NADP] cytoplasmic (EC 1.1.1.42)
 DE (Oxalouccinate decarboxylase) (IDH) (NADP+-specific ICDH) (IDP).
 GN IDH1 OR P1CD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99083434; PubMed=9866202;
 RA Nekrutenko A., Hillis D.M., Patton J.C., Bradley R.D., Baker R.J.;
 RT "Cyosolic isocitrate dehydrogenase in humans, mice, and voles and
 RT phylogenetic analysis of the enzyme family.";
 RL Mol. Biol. Evol. 15:1674-1684(1998).
 RN [2]
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
 RX MEDLINE=99452939; PubMed=10521434;
 RA Gelehrsch B.V., Gould S.J.;
 RT "The human P1CD gene encodes a cytoplasmic and peroxisomal NADP(+)-
 RT dependent isocitrate dehydrogenase.";
 RL J. Biol. Chem. 274:30527-30533(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Kidney;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wienann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Anseger W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
 RA Lander J., Duesterhoeft A., Beyer A., Koehler K., Strick N.,
 RA Mewes H.-W., Oltenswaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strunberg R.L., Feingold E.A., Gronow L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Alteschul S.F., Zeeberg B., Buelow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein W.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton B., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalski U., Smallos D.E.,
 RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 100-253 FROM N.A.
 RA Kullmann F., Vogt T., Welsh J., McClelland M.;
 RT "Differential gene expression in epithelial cells induced by bile
 RT salts: identification by RNA arbitrarily primed PCR.";
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Isocitrate + NADP(+) = 2-oxoglutarate + CO(2)
 CC + NADPH.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and peroxisomal.
 CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
 CC DEHYDROGENASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AF020038; AAD02918.1; -
 DR EMBL, AF113917; AAD9294.1; -
 DR EMBL, AL136702; CAB66637.1; -
 DR EMBL, BC012846; AAH12846.1; -
 DR EMBL, U62389; AAB17375.1; -
 DR PIR, T46280; T46280.
 DR Genew; HGNC:5382; IDH1.
 DR GK, O75874; -
 DR MIM, 147700; -
 DR GO, GO:0005829; C:cytosol; TAS.
 DR GO, GO:0004450; P:isocitrate dehydrogenase (NADP+) activity; TAS.
 DR GO, GO:0006102; P:isocitrate metabolism; TAS.
 DR InterPro, IPR001804; Isoch.
 DR InterPro, IPR004790; NADp_idh_euk.
 DR Pfam, PF00180; Isoch. 1.
 DR ProDom, PD328597; NADp_idh_euk; 1.
 DR TIGRFAMs, TIGR00127; nadp_idh_euk; 1.
 DR PROSITE, PS00470; IDH IMDH; 1-
 KW Oxidoreductase; NADP; Glyoxylate bypass; Tricarboxylic acid cycle;
 KM Peroxisome.
 FT ACT_SITE 94 94 BINDING TO ISOCITRATE (BY SIMILARITY).
 FT CONFLICT 32 32 F -> I (IN REF. 3).
 FT CONFLICT 126 126 K -> E (IN REF. 3).
 FT CONFLICT 218 218 K -> I (IN REF. 1).
 FT CONFLICT 329 329 P -> L (IN REF. 1).
 FT CONFLICT 381 381 K -> R (IN REF. 1).
 SQ SEQUENCE 414 AA; 46659 MW; 60428B0B5E5851DC CRC64;
 Query Match 42.7%; Score 38; DB 1; Length 414;
 Best Local Similarity 58.3%; Pred. No. 50;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TYPENKGTQQT 12
 DB 155 TYTPEDGTQKVT 166
 RESULT 26
 ID CH60_ACTAC STANDARD; PRT; 546 AA.
 AC P46398;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (GroEL protein).
 GN GROEL OR GROEL OR MOXA.
 OS Actinobacillus actinomycetemcomitans (Haemophilus
 OS actinomycetemcomitans).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=714;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39.
 RC STRAIN=Y4;
 RX MEDLINE=96017061; PubMed=7567064;
 RA Nakano T., Inai Y., Yamashita Y., Kusuzaki-Nagira T., Nagaoka S.,
 RA Okanashi N., Koga T., Nishihara T.;
 RT "Molecular and immunological characterization of a 64-kDa protein of
 RT Actinobacillus actinomycetemcomitans.";
 RL Oral Microbiol. Immunol. 10:151-159(1995).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 CC proper assembly of unfolded polypeptides generated under stress
 CC conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 CC 7 subunits (By similarity).

RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kertlage A.R., Gram D.E., Kyrtides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Klotzel E.F., Dougherty B.A., McEneny K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T.D., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow F.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.C.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon *Archaeoglobus fulgidus*.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000990; AAB89628.1; -.
DR PIR: H69452; H69452.
DR TIGR: AF1625; -.
DR Hypothetical protein; signal; complete proteome.
KW Hypothetical protein; signal; complete proteome.
FT SIGNAL 1 28
FT CHAIN 29 581
SQ SEQUENCE 581 AA; 64858 MW; EABA65AF4E24EB75C CRC64;

Query Match 42.7%; Score 38; DB 1; Length 581;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TYFPKGTQGYTDQI 16
Db 561 TWKGRDSQIYTDYE 576

RESULT 29
EFG_CHLMU STANDARD; PRT; 694 AA.
AC QPPTV6;
ID 16-OCT-2001 (Rel. 40, Created)
RX MEDLINE=20150255; PubMed=10684935;
RA Reed T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Kouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson M., Desoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequence of *Chlamydia trachomatis* Mopn and *Chlamydia*
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
RIBOSOME.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE002340; AAF39532.1; -.
DR PIR: G81672; G81672.
DR HSP: P13511; IDAR.
DR TIGR: TC0721; -.
DR HAMAP: MF_00054; -1.
DR InterPro: IPR004540; EF-G.
DR InterPro: IPR000795; EF-GTPbind.
DR InterPro: IPR000640; EFG_C.
DR InterPro: IPR005517; EFG_IV.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00679; EFG_C; 1.
DR Pfam: PF03764; EFG_IV; 1.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATNFACT.
DR TIGRPFAM: TIGR00484; EF-G; 1.
DR TIGRPFAM: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFRACTOR_GTP; 1.
KW Elongation factor; protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP BIND 18 25
FT NP BIND 82 86
FT NP BIND 136 139
SQ SEQUENCE 694 AA; 76521 MW; 02A6B0E17DEA319D CRC64;

Query Match 42.7%; Score 38; DB 1; Length 694;
Best Local Similarity 61.5%; Pred. No. 89;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 PPKKGTQGYTDQI 15
Db 268 FKNKGVOQLDVI 280

RESULT 30
EFG_CHLTR STANDARD; PRT; 694 AA.
AC Q84444;
ID 30-MAY-2000 (Rel. 39, Created)
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kallman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatunov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
RIBOSOME.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE001317; AAC68036.1; -
 DR PIR; F71514; F71514.
 DR HSSP; P13551; IDAR.
 DR PHC1-2DPAGE; 084444; -
 DR HAMAP; MF_00054; -; 1.
 DR InterPro; IPR004540; EF-G.
 DR InterPro; IPR000795; EF-Gpbind.
 DR InterPro; IPR000640; EF-G.
 DR InterPro; IPR005517; EF-G_IV.
 DR InterPro; IPR004161; EF-TU_D2.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00679; EF-G_C; 1.
 DR Pfam; PF03764; EF-G_IV; 1.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR PRINTS; PR00315; ELONGATNCT.
 DR TIGRFAMs; TIGR00484; EF-G; 1.
 DR TIGRFAMs; TIGR00231; small GTP; 1.
 DR PROSITE; PS00301; EF-FACTOR GTP; 1.
 DR Elongation factor; Protein biosynthesis; GTP-binding;
 KW Complete proteome.
 FT NP_BIND 18 25 GTP (BY SIMILARITY).
 FT NP_BIND 82 86 GTP (BY SIMILARITY).
 FT NP_BIND 136 139 GTP (BY SIMILARITY).
 SQ SEQUENCE 694 AA; 76537 MW; 58BDAY7CA960DAE8 CRC64;

Query Match 42.7%; Score 38; DB 1; Length 694;
 Best Local Similarity 61.5%; Pred. No. 89;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 FPKGTQGYTDQI 15
 Db 268 FPKGTQGYTDQI 280

RESULT 31
 CAN1_PIG STANDARD; PRT; 714 AA.
 AC P35750; Q29600; Q29606;
 01-JUN-1994 (Rel. 29, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 1, large (catalytic) subunit (EC 3.4.22.17) (Calcium-activated
 DE neutral proteinase) (CANP) (Mu-type) (mucanp) (micromolar-calpain).
 GN CAPN1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith T.P., Simmen F.A., Vallet J.A.;
 RT "Rapid Communication: nucleotide sequences of two isoforms of porcine
 RT micromolar calcium-activated neutral protease 1 (mucalpain) cDNA";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 326-415 FROM N.A.
 RC TISSUE=Small intestine;
 RA Winzer A.K., Fredholm M., Davies W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 528-623 FROM N.A.
 RC TISSUE=skeletal muscle;
 RA MEDLINE=9414615; PubMed=8312396;
 RA Sun W., Ji S.Q., Ebert P.J., Bidwell C.A., Hancock D.L.;

RT "Cloning the partial cDNAs of mu-calpain and m-calpain from porcine
 RT skeletal muscle".
 RL Biochimie 75:931-936(1993).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-[Xaa, Met]-Xaa or
 CC Arg-[Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions (By similarity).
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding (By similarity).
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF263610; AAF73444.1; -
 DR EMBL; F14611; CA23154.1; -
 DR EMBL; U01180; AA65125.1; -
 DR HSSP; Q07009; IDPO.
 DR MEROPS; C02.001; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR001691; Shpoc acase.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ethand; 3.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cybpc; 1.
 DR SMART; SM00054; Eph; 3.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KW Hydroxylase; Thiol protease; Calcium-binding; Multigene family.
 FT DOMAIN 30 365 CALPAIN.
 FT DOMAIN 366 526
 FT DOMAIN 527 542 DOMAIN III.
 FT DOMAIN 543 713 LINKER.
 FT CA_BIND 554 565 DOMAIN IV.
 FT CA_BIND 566 598 EF-HAND 1.
 FT CA_BIND 628 639 EF-HAND 2.
 FT CA_BIND 663 674 EF-HAND 3.
 FT DOMAIN 693 704 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT ACT_SITE 115 115 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 FT ACT_SITE 272 272 BY SIMILARITY.
 FT ACT_SITE 296 296 BY SIMILARITY.
 FT CONFLICT 528 528 V -> I (IN REF. 3).
 FT CONFLICT 531 531 I -> N (IN REF. 3).
 FT CONFLICT 541 541 E -> G (IN REF. 3).
 FT CONFLICT 622 622 S -> A (IN REF. 3).
 SQ SEQUENCE 714 AA; 81738 MW; 0BB31DE4FC5363A CRC64;

Query Match 42.7%; Score 38; DB 1; Length 714;
 Best Local Similarity 30.0%; Pred. No. 91;
 Matches 9; Conservative 4; Mismatches 3; Indels 14; Gaps 1;

Qy 1 TYPK-----GTQGYTDQIE 16
 Db 500 TFPKKGDFVLPFPSEKAGTQIEDVDQ 529

RESULT 32
 GAC_BREDI STANDARD; PRT; 720 AA.
 AC 0915D6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutaryl 7-aminocapthaloporanic acid acylase precursor (EC 3.5.1.-)
 OS (Glutaryl-7-ACA-acylase).
 DE Brevundimonas diminuta (Pseudomonas diminuta).
 OS Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Brevundimonas.
 NCBI_TaxID=293;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KAC-1;
 RA Kim D.-W., Kang S.-M., Yoon K.-H.;
 RT "Cloning and the nucleotide sequence of a Pseudomonas diminuta KAC-1
 glutaryl 7-aminocapthaloporanic acid acylase gene";
 RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 -1- PATHWAY: SECOND STEP IN THE BIOCONVERSION OF CEPHALOSPORIN TO 7-
 ACA.
 -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S45.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF51710; AAF64242.1; -
 CC PDB: 1FM2; 15-AUG-01.
 CC PDB: 1JW2; 01-SEP-02.
 CC PDB: 1JW0; 01-SEP-02.
 CC PDB: 1KEH; 04-DEC-02.
 CC MEROPS: S45.002; -
 CC InterPro: IPR002692; Penicillinamidase.
 CC DR Pfam: PF01804; Penicillinamidase; 1.
 CC KW Hydrolase; Periplasmic; Antibiotic resistance; Zymogen; Signal;
 CC 3D-structure.
 CC FT SIGNAL 1 27 POTENTIAL.
 CC FT CHAIN 28 720 GLUTARYL 7-AMINOCAPHALOSPORANIC ACID
 CC FT ACT SITE 199 199 ACYLASE.
 CC FT SEQUENCE 720 AA; 79779 MW; AD624797845CC39B CRC64;
 CC SO PROBABLE.
 CC Query Match 42.7%; Score 38; DB 1; Length 720;
 CC Best Local Similarity 70.0%; Pred. No. 92;
 CC Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 7 GTQGYTDQIE 16
 DB 680 GTTHYSQDIE 689
 RESULT 33
 SSYL_YEAST STANDARD; PRT; 852 AA.
 ID SSYL_YEAST
 AC 003770;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Putative amino-acid permease SSY1.
 OS SSYL OR SHR10 OR YDR160W OR YD8358.14.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 CC [1]
 RN SEQUENCE FROM N.A.

RC STRAIN=S288c / AB972;
 RA Murphy L., Richards C., Harrie D., Barrell B.G., Rajandream M.A.,
 RA Walsh S.V.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP IDENTIFICATION.
 RX MEDLINE=98144794; PubMed=9483800;
 RA Joergensen M.V., Bruun M.B., Didion T., Kjelgaard-Brandt M.C.;
 RT "Mutations in five loci affecting GAB1-independent uptake of neutral
 RT amino acids in yeast";
 RT Yeast 14:103-114(1998).
 CC -1- FUNCTION: MAY BE INVOLVED IN AMINO ACID UPTAKE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).
 CC -1- SIMILARITY: Belongs to the amino acid permease family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z50046; CAA90380.1; -
 CC PIR: S57984; S57984.
 CC DR SGD: S0002567; SSYL.
 CC DR GO: GO:0005886; Cytoplasmic membrane; IDA.
 CC DR GO: GO:0016597; P:amino acid binding activity; IGI.
 CC DR GO: GO:0007606; P:chemosensory perception; IGI.
 CC DR InterPro: IPR002293; AA/rel_permease1.
 CC DR InterPro: IPR004840; AAC permease.
 CC DR InterPro: IPR004841; Permease.
 CC DR InterPro: IPR004762; Yeast AA perm.
 CC DR Pfam: PF00324; aa_permease; 1.
 CC DR TIGRfams: TIGR00913; ZAO310; 1.
 CC KW PROSITE; PS00218; AMINO ACID PERMEASE 1; FALSE NEG.
 CC DR Transport; Amino-acid transport; Transmembrane.
 CC FT TRANSMEM 286 306 POTENTIAL.
 CC FT TRANSMEM 307 327 POTENTIAL.
 CC FT TRANSMEM 330 350 POTENTIAL.
 CC FT TRANSMEM 358 378 POTENTIAL.
 CC FT TRANSMEM 401 421 POTENTIAL.
 CC FT TRANSMEM 423 443 POTENTIAL.
 CC FT TRANSMEM 501 521 POTENTIAL.
 CC FT TRANSMEM 541 561 POTENTIAL.
 CC FT TRANSMEM 624 644 POTENTIAL.
 CC FT TRANSMEM 674 694 POTENTIAL.
 CC FT TRANSMEM 704 724 POTENTIAL.
 CC FT TRANSMEM 756 776 POTENTIAL.
 CC FT TRANSMEM 785 805 POTENTIAL.
 CC SO SEQUENCE 852 AA; 95744 MW; 81A94141D94EEF7D CRC64;
 CC Query Match 42.7%; Score 38; DB 1; Length 852;
 CC Best Local Similarity 63.6%; Pred. No. 11e+02;
 CC Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 3 FENKGTQGYTD 13
 DB 11 FENKNIQFTD 21
 RESULT 34
 HYSA_STRPN STANDARD; PRT; 1066 AA.
 ID HYSA_STRPN
 AC Q54873; Q54874;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hyaluronate lyase precursor (EC 4.2.2.1) (Hyaluronidase) (HYase).
 OS SP0314.
 OS Streptococcus pneumoniae.
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.

OX NCBI_TaxId=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RA MEDLINE=21357209; PubMed=11463916;
 RA Tetteilin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Haidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gilm M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Rauden D.,
 RA Holtapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angluon S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of *Streptococcus*
 RT *pneumoniae*."; Science 293:498-506(2001).
 RL [2]
 RN SEQUENCE OF 89-1066 FROM N.A.
 RP STRAIN=Type 23;
 RA MEDLINE=94156460; PubMed=8112843;
 RA Berry A.M., Lock R.A., Thomas S.M., Rajan D.P., Hansman D.,
 RA Paton J.C.;
 RT "Cloning and nucleotide sequence of the *Streptococcus pneumoniae*
 RT hyaluronidase gene and purification of the enzyme from recombinant
 RT *Escherichia coli*."; Infect. Immun. 62:1101-1108(1994).
 RL [3]
 RN CRYSTALLIZATION
 RP MEDLINE=98234706; PubMed=9573623;
 RA Jeddzejcs W.J., Chantelat L., Mewbourne R.B.;
 RT "Crystallization and preliminary X-ray analysis of *Streptococcus*
 RT *pneumoniae* hyaluronate lyase."; J. Struct. Biol. 121:73-75(1998).
 RL [4]
 RN -1- CATALYTIC ACTIVITY: Hyaluronate = N-3-(4-deoxy-beta-D-gluc-4-
 RN enuronosyl)-N-acetyl-D-glucosamine.
 RL -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan
 RL CC
 RL -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 8.
 RL CC
 RL This SWISS-PROT entry is copyright. It is produced through a collaboration
 RL between the Swiss Institute of Bioinformatics and the EMBL outstation-
 RL CC the European Bioinformatics Institute. There are no restrictions on its
 RL use by non-profit institutions as long as its content is in no way
 RL modified and this statement is not removed. Usage by and for commercial
 RL entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 RL or send an email to license@isb-sib.ch).
 RL CC
 RL EMBL; AE007344; AAK74491.1; -
 RL EMBL; L20670; AAK53685.1; ALT_INIT.
 RL EMBL; L20670; AAK53686.1; ALT_INIT.
 RL PIR; B95037; B95037.
 RL PDB; 1C82; 05-APR-01.
 RL PDB; 1P9G; 16-MAY-01.
 RL PDB; 1LOH; 07-AUG-02.
 RL PDB; 1LXK; 07-AUG-02.
 RL PDB; 1N7N; 31-DEC-02.
 RL PDB; 1N7O; 31-DEC-02.
 RL PDB; 1N7P; 31-DEC-02.
 RL PDB; 1N7Q; 31-DEC-02.
 RL PDB; 1N7R; 31-DEC-02.
 RL TIGR; SP0314; -
 RL InterPro; IPR001899; Gram_pos_anchor.
 RL InterPro; IPR006192; LPXNG.
 RL InterPro; IPR003159; Lyase_8.
 RL InterPro; IPR004103; Lyase_8_C.
 RL Pfam; PF02018; CBM_4_9; 1.
 RL Pfam; PF00746; Gram_pos_anchor; 1.
 RL Pfam; PF02278; Lyase_8; 1.
 RL Pfam; PF02884; Lyase_8_C; 1.
 RL TIGR; TIGR01167; LPXNG_anchor; 1.
 RL PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
 DR DR Lyase; Cell wall; Peptidoglycan-anchor; Signal; Complete proteome;
 KW 3D-structure.

FT SIGNAL 1 30
 FT CHAIN 31 1039
 FT PROPER 1040 1066
 FT SITE 1036 1040
 FT MOD_RES 1039 1039
 FT CONFLICT 108 108
 FT CONFLICT 115 115
 FT CONFLICT 139 139
 FT CONFLICT 211 211
 FT CONFLICT 226 226
 FT CONFLICT 230 230
 FT CONFLICT 313 313
 FT CONFLICT 340 340
 FT CONFLICT 613 613
 FT CONFLICT 658 658
 FT CONFLICT 821 821
 FT CONFLICT 848 848
 FT CONFLICT 853 853
 FT CONFLICT 907 907
 SQ SEQUENCE 1066 AA; 120771 MW; 81DB2A837BE61F9 CRC64;
 Query Match 42.7%; Score 38; DB 1; Length 1066;
 Best Local Similarity 46.7%; Pred. No. 1.4e+02;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 YPNNKTOOYDQIE 16
 Db 429 YFSDEIKKTYDVIE 443
 RESULT 35
 LAR CAEEL STANDARD; PRT: 2200 AA.
 AC Q9BMN8; Q09434; Q17859; Q20137; Q9BMN7;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein-tyrosine phosphatase Lar-like precursor (EC 3.1.3.48)
 DE (Protein-tyrosine phosphatase phosphohydrolase ptp-3).
 GN PTP-3 OR C09D8.1/C09D8.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE SPLICING.
 RC STRAIN=Bristol N2;
 RA MEDLINE=21956343; PubMed=11959824;
 RA Harrington R.J., Gutch M.J., Hengartner M.O., Tonks N.K.,
 RA Chisholm A.D.;
 RT "The C. elegans LAR-like receptor tyrosine phosphatase PTP-3 and the
 RT VAB-1 Eph receptor tyrosine kinase have partly redundant functions in
 RT morphogenesis."; Development 129:2141-2153 (2002).
 RL [2]
 RL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN=Bristol N2;
 RA Coles L., Swindurne J.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Has a role in early neural and epidermal development;
 CC neuroblast movements during closure of the gastrulation cleft and
 CC epidermal morphogenesis. Vab-1 and ptp-3 may function redundantly
 CC within the same sets of neuronal precursors.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein; at adherens
 CC junctions.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;

RL J. Biol. Chem. 269:32992-32998(1994).
 RN [4]
 RP SEQUENCE OF 2711-3396 FROM N.A.
 RC TISSUE=Lung fibroblast;
 RX MEDLINE=88007514; PubMed=2820964;
 RA Krusius T., Gehlsen K.R., Ruoslahti E.;
 RT A fibroblast chondroitin sulfate proteoglycan core protein contains
 RT lectin-like and growth factor-like sequences.";
 RL J. Biol. Chem. 262:13120-13125(1987).
 RN [5]
 RP SEQUENCE OF 251-347 FROM N.A.
 RX MEDLINE=93122792; PubMed=1478664;
 RA Iozzo R.V., Naoi M.F., Cannizzaro L.A., Wasmuth J.J.,
 RA McPherson J.D.;
 RT Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
 RT human chromosome 5 (5q12-5q14)."
 RL Genomics 14:845-851(1992).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM V3).
 RX TISSUE=Brain;
 MEDLINE=95181355; PubMed=7876137;
 RA Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
 RT Expression of PG-M(V3), an alternatively spliced form of PG-M
 RT without a chondroitin sulfate attachment in region in mouse and human
 RT tissues.";
 RL J. Biol. Chem. 270:3914-3918(1995).
 RN [7]
 RP SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
 RX TISSUE=Aortic smooth muscle;
 MEDLINE=99327053; PubMed=10397680;
 RA Lemire J.M., Braun K.R., Maupel P., Kaplan E.D., Schwartz S.M.,
 RA Wright T.N.;
 RT Versican/PG-M isoforms in vascular smooth muscle cells.";
 RT Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
 RL [8]
 RP PARTIAL SEQUENCE.
 RX TISSUE=Brain;
 MEDLINE=89174663; PubMed=2466833;
 RA Perides G., Lane W.S., Andrews D., Dahl D., Biglami A.;
 RT Isolation and partial characterization of a glial
 RT hyaluronate-binding protein.";
 RL J. Biol. Chem. 264:5981-5987(1989).
 RN [9]
 RP TISSUE SPECIFICITY OF ISOFORMS.
 MEDLINE=96213482; PubMed=8627343;
 RA Paulus W., Baur I., Douze-Zimmermann M.T., Zimmermann D.R.;
 RX "Differential expression of versican isoforms in brain tumors.";
 J. Neurochem. Exp. Neurol. 55:528-533(1996).
 RL -1- FUNCTION: May play a role in intercellular signaling and in
 RL connecting cells with the extracellular matrix. May take part in
 RL the regulation of cell motility, growth and differentiation. Binds
 RL hyaluronic acid.
 CC -1- SUBUNIT: Interacts with FBLN1 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing: Named isoforms=5;
 CC Comment=Additional isoforms seem to exist;
 CC Name=V0;
 CC IsoId=PI3611-1; Sequence=Displayed;
 CC Name=V1;
 CC IsoId=PI3611-2; Sequence=VSP_003082, VSP_003083;
 CC Name=V2;
 CC IsoId=PI3611-3; Sequence=VSP_003084;
 CC Name=V3;
 CC IsoId=PI3611-4; Sequence=VSP_003082, VSP_003085;
 CC Name=Vint;
 CC IsoId=PI3611-5; Sequence=VSP_003086;
 CC -1- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
 CC in normal brain, gliomas, medulloblastomas, schwannomas,
 CC neurofibromas, and meningiomas; v2 is restricted to normal brain
 CC and gliomas; v3 is found in all these tissues except
 CC medulloblastomas.
 CC -1- DEVELOPMENTAL STAGE: Disappears after the cartilage development.

CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- SIMILARITY: Contains 2 link domains.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -1- SIMILARITY: Contains 1 Sushi (SCR) domain.
 CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, U16306; AAA5018.1; -
 CC EMBL, X15998; CAA54128.1; -
 CC EMBL, S52468; AAB24878.1; -
 CC EMBL, U26555; AAA67565.1; -
 CC EMBL, D32039; BAA06801.1; -
 CC EMBL, J02814; AAA36437.1; -
 CC EMBL, AF084545; AAD48545.1; -
 CC FIR, S06014; A60979.
 CC HSSP, P01132; IEGF.
 CC Genew; HGNC:2464; CSPG2.
 CC MIM, 118661; -
 CC GO, GO:0005578; C:extracellular matrix; TAS.
 CC GO, GO:0005204; C:chondroitin sulfate proteoglycan; TAS.
 CC GO, GO:0005540; C:hyaluronic acid binding activity; TAS.
 CC GO, GO:0008037; P:cell recognition; TAS.
 CC GO, GO:007275; P:development; TAS.
 CC InterPro, IPR000152; Asx hydroxyl.
 CC InterPro, IPR000742; EGF_2.
 CC InterPro, IPR001881; EGF_Ca.
 CC InterPro, IPR006209; EGF_Like.
 CC InterPro, IPR007110; Ig-Like.
 CC InterPro, IPR003599; Ig.
 CC InterPro, IPR003006; Ig_MHC.
 CC InterPro, IPR001304; Lectin_C.
 CC InterPro, IPR000538; Link.
 CC InterPro, IPR000436; Sushi_SCR_CCP.
 CC Pfam, PF00008; EGF_2.
 CC Pfam, PF00059; lectin_C_1.
 CC Pfam, PF00084; sushi_1.
 CC Pfam, PF00193; Xlink_2.
 CC ProDom, PD000918; Link_2.
 CC SMART, SM00032; CCP_1.
 CC SMART, SM00034; CLECT_1.
 CC SMART, SM00179; EGF_CA_1.
 CC SMART, SM00409; IG_1.
 CC SMART, SM00445; LINK_2.
 CC PROSITE, PS00010; ASX_HYDROXYL_1.
 CC PROSITE, PS00615; C_TYPE_LECTIN_1_1.
 CC PROSITE, PS50041; C_TYPE_LECTIN_2_1.
 CC PROSITE, PS00022; EGF_1_2.
 CC PROSITE, PS01186; EGF_2_1.
 CC PROSITE, PS01187; EGF_CA_1.
 CC PROSITE, PS50835; IG_LIKE_1.
 CC PROSITE, PS01241; LINK_2.
 CC GlycoProfile; Proteoglycan; Lectin; Extracellular matrix; Sushi;
 CC Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 CC Hyaluronic acid; Alternative splicing.
 CC SIGNAL 1 20 POTENTIAL.
 CC CHAIN 21 3396 VERSICAN CORE PROTEIN.
 CC DOMAIN 21 146 IG-LIKE V-TYPE.
 CC DOMAIN 167 244 LINK 1.
 CC DOMAIN 265 346 LINK 2.
 CC DOMAIN 348 1335 GAG-ALPHA
 CC (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
 CC FT 1336 3089 GAG-BETA.
 CC FT 3089 3125 EGF-LIKE 1.
 CC FT 3127 3163 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 CC FT DOMAIN 3176 3290 C-TYPE LECTIN.

FT DOMAIN 3295 3353 SUSHI.
 FT DISULFID 44 130 BY SIMILARITY.
 FT DISULFID 172 243 BY SIMILARITY.
 FT DISULFID 196 217 BY SIMILARITY.
 FT DISULFID 270 345 BY SIMILARITY.
 FT DISULFID 294 315 BY SIMILARITY.
 FT DISULFID 3093 3104 BY SIMILARITY.
 FT DISULFID 3113 3124 BY SIMILARITY.
 FT DISULFID 3131 3142 BY SIMILARITY.
 FT DISULFID 3136 3151 BY SIMILARITY.
 FT DISULFID 3153 3162 BY SIMILARITY.
 FT DISULFID 3169 3180 BY SIMILARITY.
 FT DISULFID 3197 3289 BY SIMILARITY.
 FT DISULFID 3265 3281 BY SIMILARITY.
 FT DISULFID 3296 3339 BY SIMILARITY.
 FT DISULFID 3325 3352 BY SIMILARITY.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Query Match
 Best Local Similarity 42.7%; Score 38; DB 1; Length 3396;
 Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 1 TYFPNKGTOQYTDQIE 16
 689 TLIPERMRTDTYDEIQ 704
 RESULT 37
 NTOX RICPR STANDARD; PRT; 110 AA.
 AC 09ZGZ2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADH-quinone oxidoreductase chain K (EC 1.6.99.5) (NADH dehydrogenase
 I, chain K) (NDH-1, chain K).
 GN NTOX OR Rp791.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Medid B;
 RC MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 Sickeritz-Ponten T., Alemark U.C.M., Podowski R.M., Naslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria."
 RL Nature 396:133-140 (1998).
 CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
 sulfur (Fe-S) centers, to quinones in the respiratory chain.
 CC Couples the redox reaction to proton translocation (for every two
 electrons transferred, four hydrogen ions are translocated across
 the cytoplasmic membrane), and thus conserves the redox energy in
 a proton gradient (by similarity).
 CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO POLYPEPTIDE 4L OF THE NADH-UBIQUINOL OXIDOREDUCTASE
 OF CHLOROPLASTS OR MITOCHONDRIA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; AJ235273; CAA15217.1; -.
 DR PIR; A71640; A71640.
 DR InterPro; IPR003215; NADH_dh_ubiq1.
 DR InterPro; IPR001133; Oxidored_4L.
 DR InterPro; IPR003214; Oxidored_4L.
 DR Pfam; PF00420; Oxidored_q2; 1.
 DR ProDom; PD002107; NADH_dh_ubiq1; 1.
 KW Oxidoreductase; NAD; Quinone; Transmembrane; Complete proteome.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 SQ SEQUENCE 110 AA; 12423 MW; F62159A6410FA35D CRC64;
 Query Match
 Best Local Similarity 41.6%; Score 37; DB 1; Length 110;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 2 YFPNKGTOQYTD 13
 93 YFPNKGSIKTD 104
 Db 93 YFPNKGSIKTD 104
 RESULT 38
 YAS1_ARCFU STANDARD; PRT; 161 AA.
 AC 029211;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1051 precursor.
 GN AF1051.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2254;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RC MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kellavagge A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus."
 RL Nature 390:364-370 (1997).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

QY 2 YFPNKGNOOYTD 13
DB 74 YFNTGDERYPD 85

RESULT 39
RS9_STRCO STANDARD; PRT; 170 AA.
ID RS9_STRCO
AC 053875;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S9.
GN RPS1 OR SC04735 OR SC6G4.13.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
[1]

SEQUENCE FROM N.A.
STRAIN=3(2);
MEDLINE=98102746; PubMed=9439573;
Sanchez C., Blanco G., Mendez C., Salas J.A.;
"Cloning, sequencing and transcriptional analysis of a Streptomyces
coelicolor operon containing the rplM and rplI genes encoding
ribosomal proteins Scoli3 and Scoli5".
Mol. Gen. Genet. 257:91-96 (1997).
[2]

SEQUENCE FROM N.A.
STRAIN=3(2) / M145;
MEDLINE=21996410; PubMed=12000953;
Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleier H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,
Huang C.-H., Kleier T., Larke L., Murphy L., Oliver K., O'Neil S.,
Rabbinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wierzbicki A., Woodward J., Barrall B.G., Parkhill J.,
Hopwood D.A.;
"Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)".
Nature 417:141-147 (2002).
[1]

-1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; U43429; AAC46061.1; -
DR EMBL; AL939121; CAA20391.1; -
CC PIR; T35564; T35564.
DR HAMAP; MF_00532; -
DR InterPro; IPR000754; RIBOSOMAL_S9.
DR Pfam; PFO0380; RIBOSOMAL_S9; 1.
DR ProDom; PD001627; RIBOSOMAL_S9; 1.
DR PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 170 AA; 18699 MW; 3CE80DF69151C99 CRC64;

Query Match 41.6%; Score 37; DB 1; Length 170;
Best Local Similarity 53.8%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YFPNKGNOOYTD 14
DB 77 YFPNKGNOOEVNE 89

RESULT 40
PACA_ONCNE STANDARD; PRT; 173 AA.
ID PACA_ONCNE
AC P41585;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucagon-family neuropeptides precursor [Containing: Growth hormone-
releasing factor (GRF) (Growth hormone-releasing hormone) (GHRH);
Pituitary adenylate cyclase activating polypeptide (PACAP)].
OS Oncorhynchus nerka (Sockeye salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8023;
[1]

SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP TISSUE=Brain;
RC MEDLINE=93345532; PubMed=8344311;
Parker D.B., Coe I.R., Dixon G.H., Sherwood N.M.;
"Two salmon neuropeptides encoded by one brain cDNA are structurally
related to members of the glucagon superfamily.";
Eur. J. Biochem. 215:439-448 (1993).
RL

-1- FUNCTION: PRIMARY ROLE OF GHRH IS TO RELEASE GH FROM THE
PITUITARY.
CC

-1- FUNCTION: PACAP PLAYS PIVOTAL ROLES AS A NEUROTRANSMITTER AND/OR A
NEUROMODULATOR.
CC

-1- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P41585-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P41585-2; Sequence=VSP_001762, VSP_001763;
CC Note=Lacks the GHRH-like sequence;
CC -1- POLYMORPHISM: FOUR CLONES WERE IDENTIFIED THAT HAD NUCLEOTIDE
DIFFERENCES.
CC

-1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X73233; CAA51705.1; ALT_SEQ.
DR PIR; S34767; S34767.
DR InterPro; IPR000532; GLUCAGON.
DR Pfam; PFO0123; hormone2; 2.
DR PRINTS; PR00275; GLUCAGON.
DR SMART; SM00070; GLUCA; 2.
DR PROSITE; PS00260; GLUCAGON; 2.
KW Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW Amidation; Alternative splicing; Polymorphism.
FT SIGNAL. 1 22
FT PROPEP 23 80
FT PEPTIDE 82 126
FT PEPTIDE 129 166
FT MOD_RES 166 166
FT VARSPLIC 78 78
FT VARSPLIC 79 113
FT VARSPLIC 22 22
FT VARIANT 61 61
FT VARIANT 78 78
FT VARIANT 122 122
FT VARIANT 165 165
FT VARIANT 171 171

GROWTH HORMONE-RELEASING FACTOR.
PITUITARY ADENYLATE CYCLASE ACTIVATING
POLYPEPTIDE.
AMIDATION (G-167 PROVIDE AMIDE GROUP)
(POTENTIAL).
G->S (in isoform Short).
/FTId=VSP_001762.
Missing (in isoform Short).
/FTId=VSP_001763.
S->C.
P->S.
G->R.
T->S.
N->S.
G->A.

SQL SEQUENCE 173 AA; 19704 MM; 2B0B554F43C738F2 CRC64;

Query Match 41.6%; Score 37; DB 1; Length 173;

Best Local Similarity 50.0%; Pred. No. 29;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 YFPNKGTOOYTD 13

DB 73 YPPEKGTERRHAD 84

Search completed: August 20, 2003, 09:30:39
Job time : 11.2371 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 46.5155 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-3
Perfect score: 89
Sequence: 1 TYFPNKGTYQYTDQIE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTREMBL_23:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_mammal:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	138	12	085155 human parvo
2	89	100.0	141	12	085171 human parvo
3	89	100.0	141	12	085166 human parvo
4	89	100.0	141	12	085146 human parvo
5	89	100.0	142	12	085168 human parvo
6	89	100.0	142	12	085173 human parvo
7	89	100.0	144	12	085138 human parvo
8	89	100.0	145	12	085181 human parvo
9	89	100.0	145	12	085161 human parvo
10	89	100.0	146	12	085158 human parvo
11	89	100.0	147	12	085142 human parvo
12	89	100.0	148	12	085150 human parvo
13	89	100.0	151	12	085131 human parvo
14	89	100.0	151	12	085177 human parvo
15	89	100.0	153	12	085123 human parvo
16	89	100.0	157	12	085196 human parvo

17	89	100.0	162	12	085135 human parvo
18	89	100.0	546	12	0913X0 human parvo
19	89	100.0	546	12	0913W7 human parvo
20	89	100.0	554	12	09PZS9 human parvo
21	89	100.0	554	12	090201 human parvo
22	89	100.0	554	12	065790 human parvo
23	89	100.0	554	12	0912B7 human eryth
24	89	100.0	554	12	09JGP7 human eryth
25	89	100.0	554	12	09WKL9 human parvo
26	89	100.0	554	12	08JNS4 human parvo
27	89	100.0	760	12	09PZT8 human parvo
28	89	100.0	761	12	09PZU0 human parvo
29	89	100.0	765	12	09PZT6 human parvo
30	89	100.0	769	12	09PZT4 human parvo
31	89	100.0	773	12	0913X1 human parvo
32	89	100.0	773	12	0913W8 human parvo
33	89	100.0	781	12	08JNS6 human parvo
34	89	100.0	781	12	065789 human parvo
35	89	100.0	781	12	P90223 human parvo
36	89	100.0	781	12	085191 human parvo
37	89	100.0	781	12	P90221 human parvo
38	89	100.0	781	12	P89318 human parvo
39	89	100.0	781	12	08JNS3 human parvo
40	89	100.0	781	12	P90224 human parvo
41	89	100.0	781	12	09PZT0 human parvo
42	89	100.0	781	12	P89316 human parvo
43	89	100.0	781	12	P89320 human parvo
44	89	100.0	781	12	085117 human parvo
45	89	100.0	781	12	P89321 human parvo

ALIGNMENTS

RESULT 1
ID 085155 PRELIMINARY; PRT; 138 AA.
AC 085155;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DR Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; SEDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_taxid=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanter A., Von Pohlitzki A., Giegler A., Caasinocti P., Siegl G.,
RT "XXXXSequence variability among different parvovirus B19 isolates."
RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70564; CAA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899FB8879A3E68D CRC64;
Query Match 100.0%; Score 89; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TYFPNKGTYQYTDQIE 16
DB 103 TYFPNKGTYQYTDQIE 118
RESULT 2
ID 085171 PRELIMINARY; PRT; 141 AA.
AC 085171;
DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70580; CA94493.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT SEQUENCE 141 AA; 15770 MW; C9B92572A78C6C27 CRC64;
 SQ
 Query Match 100.0%; Score 89; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 105 TYFPNKGTOQYTDQIE 120

RESULT 3

ID 085166 PRELIMINARY; PRT; 141 AA.
 AC 085166;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70575; CA94488.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT SEQUENCE 141 AA; 15785 MW; 4115D3D915751757 CRC64;
 SQ
 Query Match 100.0%; Score 89; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 105 TYFPNKGTOQYTDQIE 120

RESULT 4

ID 085146 PRELIMINARY; PRT; 141 AA.
 AC 085146;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70555; CA94467.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT SEQUENCE 141 AA; 15712 MW; 32F6B32E7E6B8583 CRC64;
 SQ
 Query Match 100.0%; Score 89; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 102 TYFPNKGTOQYTDQIE 117

RESULT 5

ID 085168 PRELIMINARY; PRT; 142 AA.
 AC 085168;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70577; CA94490.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 142
 FT SEQUENCE 142 AA; 15821 MW; 2B6E4D9A6784F8C4 CRC64;
 SQ
 Query Match 100.0%; Score 89; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 101 TYFPNKGTOQYTDQIE 116

RESULT 6

ID 085173 PRELIMINARY; PRT; 142 AA.
 AC 085173;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Gigler A., Cassinotti P., Siegl G.,

RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Z70582; CA94495.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A95E07C0BC0434 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 5.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 101 TYFPNKGTOQYTDQIE 116

RESULT 7

AC 085138 PRELIMINARY; PRT; 144 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Z70547; CA94459.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AE CRC64;

Query Match 100.0%; Score 89; DB 12; Length 144;
 Best Local Similarity 100.0%; Pred. No. 5.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYFPNKGTOQYTDQIE 16
 103 TYFPNKGTOQYTDQIE 118

RESULT 8

ID 085181 PRELIMINARY; PRT; 145 AA.
 AC 085181;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Z70590; CA94503.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.

FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1E1B9B923C4 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 5.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 105 TYFPNKGTOQYTDQIE 120

RESULT 9

ID 085161 PRELIMINARY; PRT; 145 AA.
 AC 085161;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Z70570; CA94483.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 5.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 101 TYFPNKGTOQYTDQIE 116

RESULT 10

ID 085158 PRELIMINARY; PRT; 146 AA.
 AC 085158;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Z70567; CA94480.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6E4D9A CRC64;

Query Match 100.0%; Score 89; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 5.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
DB 101 TYFPNKGTOOYTDQIE 116

RESULT 11

085142 PRELIMINARY; PRT; 147 AA.

AC 085142; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;
[1] -

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXSSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBD databases.
DR EMBL; Z70551; CA94463.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1
FT 147
SQ SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 147;
Best Local Similarity 100.0%; Pred. No. 5.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
DB 104 TYFPNKGTOOYTDQIE 119

RESULT 12

085150 PRELIMINARY; PRT; 148 AA.

AC 085150; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;
[1] -

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXSSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBD databases.
DR EMBL; Z70559; CA94471.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1
FT 148
SQ SEQUENCE 148 AA; 16539 MW; 8814ECF2459B308B CRC64;

Query Match 100.0%; Score 89; DB 12; Length 148;
Best Local Similarity 100.0%; Pred. No. 5.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
DB 104 TYFPNKGTOOYTDQIE 119

DB 104 TYFPNKGTOOYTDQIE 119

RESULT 13

085131 PRELIMINARY; PRT; 151 AA.

AC 085131; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;
[1] -

SEQUENCE FROM N.A.

RX MEDLINE=96332516; PubMed=8760426;
RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "Sequence variability among different parvovirus B19 isolates."
RL J. Gen. Virol. 77:1781-1785(1996).
DR EMBL; Z70540; CA94452.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1
FT 151
SQ SEQUENCE 151 AA; 16902 MW; 3CA74914B8E73A3E CRC64;

Query Match 100.0%; Score 89; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
DB 105 TYFPNKGTOOYTDQIE 120

RESULT 14

085177 PRELIMINARY; PRT; 151 AA.

AC 085177; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DR 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;
[1] -

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXSSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/GenBank/DBD databases.
DR EMBL; Z70586; CA94499.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON_TER 1
FT 151
SQ SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
DB 105 TYFPNKGTOOYTDQIE 120

RESULT 15

Q85123 PRELIMINARY; PRT; 153 AA.
 AC Q85123;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96332516; PubMed=8760426;
 RA Hemanuer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "Sequence variability among different parvovirus B19 isolates."
 RL J. Gen. Virol. 77:1781-1785(1996).
 DR EMBL; Z70532; CAA9444.1; -.
 InterPro; IPR001403; Parvo_coat.
 Pfam; PF00740; Parvo_coat; 1.
 NON_TER 1
 FT NON_TER 153
 SQ SEQUENCE 153 AA; 17131 MW; ECCALF44020814EC CRC64;
 Query Match 100.0%; Score 89; DB 12; Length 153;
 Best Local Similarity 100.0%; Pred. No. 6.1e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 104 TYFPNKGTOQYTDQIE 119

RESULT 16
 ID Q85196 PRELIMINARY; PRT; 157 AA.
 AC Q85196;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Hemanuer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 RX Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70603; CAA94518.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 NON_TER 1
 FT NON_TER 157
 SQ SEQUENCE 157 AA; 17591 MW; 1DA2191887FFCE03 CRC64;
 Query Match 100.0%; Score 89; DB 12; Length 157;
 Best Local Similarity 100.0%; Pred. No. 6.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 106 TYFPNKGTOQYTDQIE 121

RESULT 17
 ID Q85135 PRELIMINARY; PRT; 162 AA.
 AC Q85135;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 GN Viral protein 1 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemanuer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70544; CAA94456.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 NON_TER 1
 FT NON_TER 162
 SQ SEQUENCE 162 AA; 18005 MW; 3F51443566660F2 CRC64;
 Query Match 100.0%; Score 89; DB 12; Length 162;
 Best Local Similarity 100.0%; Pred. No. 6.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 105 TYFPNKGTOQYTDQIE 120

RESULT 18
 ID Q913X0 PRELIMINARY; PRT; 546 AA.
 AC Q913X0;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Lali;
 RA Hokykar K., Soderlund-Venermo M., Ranki A., Kiviloto O., Partio E.K.,
 RA Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY044266; AAK95572.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 NON_TER 546
 FT NON_TER 546
 SQ SEQUENCE 546 AA; 59934 MW; 467BE8468A67282E5 CRC64;
 Query Match 100.0%; Score 89; DB 12; Length 546;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOQYTDQIE 16
 DB 391 TYFPNKGTOQYTDQIE 406

RESULT 19
 ID Q913W7 PRELIMINARY; PRT; 546 AA.
 AC Q913W7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAM;
 RA Hokyar K., Soderlund-Venermo M., Ranki A., Kivluoto O., Partio E.K., Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY044268; AAK5575.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 546 546
 SQ SEQUENCE 546 AA; 59910 MW; BBFLDD23695E3FCD CRC64;

Query Match 100.0%; Score 89; DB 12; Length 546;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYFPNKGTOOYTDQIE 16
 |||||
 391 TYFPNKGTOOYTDQIE 406

RESULT 20
 Q9PZS9 PRELIMINARY; PRT; 554 AA.
 AC Q9PZS9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE VP2 capsid protein.
 GN VP.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HV;
 RA Gallinella G., Venturoli S.;
 RT "B19 Genome Sequence and Structure Analysis."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF162273; AAD46615.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60833 MW; 7901FEBD65A697B6 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
 |||||
 DB 391 TYFPNKGTOOYTDQIE 406

RESULT 21
 Q90201 PRELIMINARY; PRT; 554 AA.
 AC Q90201;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE VP2 structural protein (Capsid protein VP2) (Fragment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 RP SEQUENCE FROM N.A.
 RC STRAIN=VEN;
 RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the

RT VP1/VP2 gene from multiple isolates."
 RL J. Gen. Virol. 77:2767-2774(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MI, and N8;
 RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
 RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RX MEDLINE=90218047; PubMed=2157807;
 RA Umene K., Nunoue T.;

RT "The genome type of human parvovirus B19 strains isolated in Japan during 1981 differs from types detected in 1986 to 1987: a correlation between genome type and prevalence."
 RT J. Gen. Virol. 71:983-986(1990).
 RL J. Gen. Virol. 71:983-986(1990).
 DR EMBL: U53595; AAB47453.1; -;
 DR EMBL: U53596; AAB47455.1; -;
 DR EMBL: U53593; AAB47449.1; -;
 DR EMBL: U53594; AAB47451.1; -;
 DR EMBL: U53597; AAB47457.1; -;
 DR EMBL: U53600; AAB47463.1; -;
 DR EMBL: U53601; AAB47465.1; -;
 DR EMBL: U53598; AAB47459.1; -;
 DR EMBL: U53599; AAB47461.1; -;
 DR EMBL: AB030693; BAA90290.1; -;
 DR EMBL: AB030673; BAA90268.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 554 554

SQ SEQUENCE 554 AA; 60853 MW; 98FE598FF20CB66F CRC64;
 Query Match 100.0%; Score 89; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
 |||||
 DB 391 TYFPNKGTOOYTDQIE 406

RESULT 22
 Q65790 PRELIMINARY; PRT; 554 AA.
 AC Q65790;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 RN NCBI_TaxID=10798;
 RP SEQUENCE FROM N.A.
 RC STRAIN=VEN;
 RA Echevarria Mayo J.E., Erdman D.D.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U13558; AAB83559.1; -;
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60798 MW; BA89F2B23BE4E24 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOYTDQIE 16
 |||||
 DB 391 TYFPNKGTOOYTDQIE 406

RESULT 23

Q912B7 PRELIMINARY; PRT; 554 AA.

AC Q912B7; 1 TYPENKGTQYTDQIE 16

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DE 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

GN VP2 protein.

OS Human erythrovirus V9.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=72197;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=V9;

RA Nguyen O.T.;

RT "Molecular cloning and sequencing of a novel human erythrovirus genome: new species beside B19 in the genus Erythrovirus."

RT Submitted (Sbf-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ249437; CAC80622.1; -

DR InterPro; IPR001403; Parvo.coat.

DR Pfam; PF00740; Parvo.coat; 1.

SQ SEQUENCE 554 AA; 60885 MW; BAF336CDAA80F336 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 554;

Best Local Similarity 100.0%; Pred. No. 2.5e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TYPENKGTQYTDQIE 16

391 TYPENKGTQYTDQIE 406

RESULT 24

Q9UGP7 PRELIMINARY; PRT; 554 AA.

AC Q9UGP7; 1 TYPENKGTQYTDQIE 16

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Capsid protein VP2.

GN VP2.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Rm;

RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;

RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."

RT Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030694; BAA90293.1; -

DR InterPro; IPR001403; Parvo.coat.

DR Pfam; PF00740; Parvo.coat; 1.

SQ SEQUENCE 554 AA; 60839 MW; 0DB958B3C73EF64 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 554;

Best Local Similarity 100.0%; Pred. No. 2.5e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TYPENKGTQYTDQIE 16

391 TYPENKGTQYTDQIE 406

RESULT 25

Q9WKL9 PRELIMINARY; PRT; 554 AA.

AC Q9WKL9; 1 TYPENKGTQYTDQIE 16

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Structural protein VP2.

GN VP2.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20014169; PubMed=10548133;

RA Hemmer A., Beckenlehner K., Wolf H., Lang B., Modrow S.;

RT "Acute parvovirus B19 infection in connection with a flare of systemic lupus erythematosus in a female patient."

RT J. Clin. Virol. 14:73-77(1999).

DR EMBL; AF13323; AAC99439.1; -

DR InterPro; IPR001403; Parvo.coat.

DR Pfam; PF00740; Parvo.coat; 1.

SQ SEQUENCE 554 AA; 60913 MW; 98FB588FF205C66F CRC64;

Query Match 100.0%; Score 89; DB 12; Length 554;

Best Local Similarity 100.0%; Pred. No. 2.5e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TYPENKGTQYTDQIE 16

391 TYPENKGTQYTDQIE 406

RESULT 26

Q8JNS4 PRELIMINARY; PRT; 554 AA.

AC Q8JNS4; 1 TYPENKGTQYTDQIE 16

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE VP2.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D91.1;

RX MEDLINE=22174902; PubMed=12186896;

RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Meritet J.F., Garbarg-Chenon A.;

RT "Genetic diversity within Human Erythroviruses: Identification of three genotypes."

RT J. Virol. 76:9124-9134(2002).

DR EMBL; AY083234; AAL91014.1; -

DR InterPro; IPR001403; Parvo.coat.

DR Pfam; PF00740; Parvo.coat; 1.

SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 554;

Best Local Similarity 100.0%; Pred. No. 2.5e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TYPENKGTQYTDQIE 16

391 TYPENKGTQYTDQIE 406

RESULT 27

Q9PZT8 PRELIMINARY; PRT; 760 AA.

AC Q9PZT8; 1 TYPENKGTQYTDQIE 16

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE VP1/2 (Fragment).

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN (1)
RX SEQUENCE FROM N.A.
RP STRAIN=Kat12;
RC MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161224; AAD45912.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON TER 760
SQ SEQUENCE 760 AA; 83403 MW; EA6BE0145E3A0E5A CRC64;

Query Match 100.0%; Score 89; DB 12; Length 760;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TYFPNKGTOOQYTDQIE 16
|||||
618 TYFPNKGTOOQYTDQIE 633

RESULT 28
Q9PZU0 PRELIMINARY; PRT; 761 AA.
AC Q9PZU0; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Kat11;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161223; AAD45910.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON TER 761
SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 761;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOQYTDQIE 16
|||||
DB 618 TYFPNKGTOOQYTDQIE 633

RESULT 29
Q9PZT6 PRELIMINARY; PRT; 765 AA.
AC Q9PZT6; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Kat13;
RA

RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161225; AAD45915.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON TER 765
SQ SEQUENCE 765 AA; 83998 MW; 89E2546086DCDBE8 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 765;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOQYTDQIE 16
|||||
DB 618 TYFPNKGTOOQYTDQIE 633

RESULT 30
Q9PZT4 PRELIMINARY; PRT; 769 AA.
AC Q9PZT4; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Kat14;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue."
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161226; AAD45917.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON TER 769
SQ SEQUENCE 769 AA; 84578 MW; 0749D46E5CA7BB68 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 769;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TYFPNKGTOOQYTDQIE 16
|||||
DB 618 TYFPNKGTOOQYTDQIE 633

RESULT 31
Q913X1 PRELIMINARY; PRT; 773 AA.
AC Q913X1; 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Minor virus capsid protein VPI (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Kat1;
RA Hokynar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,
Hedman K.;

RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY044266; AAK95571.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 773
 SQ SEQUENCE 773 AA; 85164 MW; D60CCA6F90B05378 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 773;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 DB 618 TYFPNKGTOQYTDQIE 633

RESULT 32
 O913W8

PRELIMINARY; PRT; 773 AA.

O913W8; 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Minor virus capsid protein VP1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN (1)
 RC SEQUENCE FROM N.A.
 RA STRAIN=HAM;
 RA Hokymar K., Soderlund-Venermo M., Ranki A., Kivuluoto O., Partio E.K., Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY044266; AAK95571.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 FT NON_TER 773
 SQ SEQUENCE 773 AA; 85140 MW; 2B86F3245F9CEB50 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 773;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 618 TYFPNKGTOQYTDQIE 633

RESULT 33

O8JUN56 PRELIMINARY; PRT; 781 AA.

O8JUN56; 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=D91.1;
 RC MEDLINE=22174902; PubMed=12186896;
 RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Merlet J.F., Garbarg-Chenon A.;
 RT "Genetic Diversity within Human Erythroviruses: Identification of Three Genotypes."
 RL J. Virol. 76:9124-9134(2002).
 DR EMBL: AY083234; AAL91013.1; -
 DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86142 MW; 4EB71FEDD41FC8F3 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 DB 618 TYFPNKGTOQYTDQIE 633

RESULT 34

O65789 PRELIMINARY; PRT; 781 AA.

O65789; 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN (1)
 RC SEQUENCE FROM N.A.
 RA STRAIN=VEN1;
 RA Echevarria Mayo J.E., Erdman D.D.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U31358; AAB3558.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 DB 618 TYFPNKGTOQYTDQIE 633

RESULT 35

P90223 PRELIMINARY; PRT; 781 AA.

P90223; 01-MAY-1997 (TRENBLrel. 03, Created)
 DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=USA2;
 RC MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigon E.V., Wang Q.Y., Anderson L.V.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the RT VP1/VP2 gene from multiple isolates."
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL: U38514; AAB47796.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TYFPNKGTOQYTDQIE 16
 DB 618 TYFPNKGTOQYTDQIE 633

RESULT 36

085191 PRELIMINARY; PRT; 781 AA.
 AC 085191;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ORF2 protein.
 GN ORF2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_Taxid=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemmer A., Von Poblotzki A., Giegler A., Casinotti P., Siegl G., Wolf H., Modrow S.;
 RT "XXSquence variability among different parvovirus B19 isolates."; Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 EMBL; 270599; CAA94513.1; -.
 InterPro: IPR001403; Parvo coat.
 DR Pfam: PR00740; Parvo coat; 1.
 SQ SEQUENCE 781 AA; 86026 MW; AA02577B63EDB2D CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOOYTDQIE 16
 Db 618 TYFPNKGTOOYTDQIE 633

RESULT 37

P90221 PRELIMINARY; PRT; 781 AA.
 AC P90221;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_Taxid=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=KOR2;
 MEDLINE=97081188; PubMed=8922470;
 Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the VP1/VP2 gene from multiple isolates."; J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38511; AAB47793.1; -.
 DR InterPro: IPR001403; Parvo coat.
 DR Pfam: PR00740; Parvo coat; 1.
 SQ SEQUENCE 781 AA; 86070 MW; 9FA830083F6F1357 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOOYTDQIE 16
 Db 618 TYFPNKGTOOYTDQIE 633

RESULT 38

P89318 PRELIMINARY; PRT; 781 AA.
 AC P89318;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_Taxid=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=JAP1;
 RX MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the VP1/VP2 gene from multiple isolates."; J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38509; AAB47791.1; -.
 DR InterPro: IPR001403; Parvo coat.
 DR Pfam: PR00740; Parvo coat; 1.
 SQ SEQUENCE 781 AA; 86064 MW; 3AB65CE69096339 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOOYTDQIE 16
 Db 618 TYFPNKGTOOYTDQIE 633

RESULT 39

Q8YB3 PRELIMINARY; PRT; 781 AA.
 AC Q8YB3;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE 7.5 kDa protein.
 OS Erythrovirus A6.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_Taxid=182494;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen Q.T., Wong S., Brown K.E.;
 RT "Identification and characterization of a second novel human erythrovirus variant, A6."; Virology 0:0-0(2002).
 DR EMBL; AY064475; AAL55418.1; -.
 DR InterPro: IPR001403; Parvo coat.
 DR Pfam: PR00740; Parvo coat; 1.
 SQ SEQUENCE 781 AA; 86277 MW; E73AB663E4551128 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOOYTDQIE 16
 Db 618 TYFPNKGTOOYTDQIE 633

RESULT 40

P90224 PRELIMINARY; PRT; 781 AA.
 AC P90224;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VP1 and VP2 structural protein.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NC NCB1_Taxid=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BRZ1;
 RX MEDLINE=97081188; PubMed=8922470;

W d Aug 20 12:54:13 2003

RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the
 VP1/VP2 gene from multiple isolates."
 RL J. Gen. Virol. 77:2767-2774(1996).
 DR EMBL; U38546; AAB47801.1;-.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86039 MW; C3A29EB1DB8AD378 CRC64;

Query Match 100.0%; Score 89; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. NO. 3.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TYFPNKGTOQYTDQIE 16
 |||||
 DB 618 TYFPNKGTOQYTDQIE 633

Search completed: August 20, 2003, 09:23:58
 time : 47.5155 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 71.4433 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-2

Perfect score: 118
Sequence: 1 GLNMTYFPNKGQYTDQIE 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	100.0	370	12	AA13406
2	118	100.0	543	12	AA13405
3	118	100.0	554	16	AAW08987
4	118	100.0	554	20	AAV23230
5	118	100.0	554	24	ABP57264
6	118	100.0	554	24	ABP57267
7	118	100.0	756	21	AAV17231
8	118	100.0	781	16	AAW08986
9	118	100.0	781	20	AAV23227

10	118	100.0	781	24	ABP57263	Human parvovirus B
11	118	100.0	781	24	ABP57266	Human parvovirus B
12	47	39.8	93	23	ABP57712	Zinc finger protei
13	47	39.8	305	23	ABP48499	Listeria monocytog
14	47	39.8	904	23	ABG13993	Novel human diagno
15	45	38.1	225	19	AAW98428	H. pylori GHP 409
16	45	38.1	702	22	ABG70124	Drosophila melanog
17	45	38.1	757	22	ABG19946	Novel human diagno
18	45	38.1	814	20	AAV11067	H. pylori ORF 13ae
19	45	38.1	814	20	AAV11062	H. pylori outer me
20	45	38.1	815	19	AAV11066	H. pylori ORF 13ae
21	45	38.1	815	20	AAV11761	H. pylori outer me
22	44.5	37.7	142	21	AAV1298	Human ORF1062
23	44.5	37.7	142	23	ABP07047	Human ORFX protein
24	44.5	37.7	406	22	ABP61725	Drosophila melanog
25	44	37.3	204	23	AAW48920	Human L1 retroposo
26	44	37.3	1017	23	ABP73753	Candida albicans e
27	43.5	36.9	616	23	ABP93919	Herbicideally activ
28	43	36.4	119	18	AAW21827	Rat RT1.Balpal c1
29	43	36.4	180	20	AAW06804	Peptide Seg ID No:
30	43	36.4	180	24	AAE33454	Rat RT1B betatalph
31	43	36.4	185	20	AAV06797	Betatalph/MBP-72
32	43	36.4	185	24	AAE33447	Rat RT1B betatalph
33	43	36.4	204	22	ABG02830	Novel human diagno
34	43	36.4	297	16	AAE72588	Carponin protein.
35	43	36.4	297	17	AAE94888	Carponin. Homo sa
36	43	36.4	493	22	ABP65502	Drosophila melanog
37	43	36.4	726	23	ABP77613	AMEV NTPase (AMV)
38	42.5	36.0	525	22	ABP70228	Drosophila melanog
39	42	35.6	78	21	AAW44935	Human secreted pro
40	42	35.6	115	23	ABP32719	Human ORF1692 prot
41	42	35.6	156	23	ABP89112	Human polypeptide
42	42	35.6	157	19	AAW74761	Human secreted pro
43	42	35.6	157	23	ABP95210	Human novel secret
44	42	35.6	218	21	AAE68772	Amino acid sequenc
45	42	35.6	218	22	AAE94972	Human protein sequ

ALIGNMENTS

RESULT 1
AA13406
ID AA13406 standard; Protein; 370 AA.
XX
AC AA13406;
XX
DT 24-OCT-1991 (first entry)
XX
DE Parvo virus B19 PANSE.
XX
KM Primer; PCR; PAPST; globulin.
XX
OS Synthetic.
XX
FN DE4003826-A.
XX
PD 14-AUG-1991.
XX
PE 08-FEB-1990; 90DE-4003826.
XX
PR 08-FEB-1990; 90DE-4003826.
XX
PA (MIKR-) MIKROGEN MOLEKULARB.
XX
PI Soutschek E, Motz M;
XX
DR WPI; 1991-246423/34.
XX
PT Immunologically active parvo virus B19 peptide(s) - comprising
PT capsid protein VPI or VP2 fragments, useful for antibody
PT detection or vaccination
XX

PS Claim 16; Page 10-11; 22pp; German.
 XX
 CC VP2 (AA13405) and its fragments PANSE and PAPST (AA13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AA013159-63 in pairs.
 CC See also AA13400-07 and AA13414.
 XX
 SQ Sequence 370 AA;

Query Match 100.0%; Score 118; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. NO. 2e-10; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLNMHTYFPNKGTOQYTDQIE 21
 |||||
 222 GLNMHTYFPNKGTOQYTDQIE 242

RESULT 2
 AA13405
 ID AA13405 standard; Protein; 543 AA.
 XX
 AC AA13405;

DT 24-OCT-1991 (first entry)
 XX
 DE Parvo virus B19 VP2.

XX Primer; PCR; globulin; PANSE; PAPST.
 XX
 OS Synthetic.

XX DB4003826-A.
 XX
 PD 14-AUG-1991.

XX PE 08-FEB-1990; 90DE-4003826.
 XX
 PR 08-FEB-1990; 90DE-4003826.

XX (MIKR-) MIKROGEN MOLEKULARB.
 XX
 PA
 XX

PT Soutschek E, Motz M;
 XX
 WPI; 1991-246423/34.

PT Immunologically active parvo virus B19 peptide(s) - comprising
 PT capsid protein VP1 or VP2 fragments, useful for antibody
 PT detection or vaccination
 XX
 PS Disclosure; Fig 2-6; 22pp; German.

XX VP2 and its fragments PANSE (AA13406) and PAPST (AA13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AA013159-63 in pairs.
 CC See also AA13400-07 and AA13414.
 XX
 SQ Sequence 543 AA;

Query Match 100.0%; Score 118; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. NO. 3.1e-10; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLNMHTYFPNKGTOQYTDQIE 21
 |||||
 DB 395 GLNMHTYFPNKGTOQYTDQIE 415

RESULT 3
 AA08987
 ID AA08987 standard; Protein; 554 AA.
 XX
 AC AA08987;

DT 27-FEB-1997 (first entry)
 XX
 DE Human parvovirus VP-2 protein.

XX Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KW non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KW erythblastemia; abortion; universal fetal hydrops; liver disease;
 KW haemorrhagic fever; rheumatism; detection; IgG antibody.

XX Human parvovirus.
 XX
 EN JP07147986-A.

XX 13-JUN-1995.
 XX
 PD 24-SEP-1992; 92JP-0281017.

XX 24-SEP-1992; 92JP-0281017.
 XX
 PR 24-SEP-1992; 92JP-0281017.

XX (DENK-) DENKA SEIKEN KK.
 XX
 PA (ELED) KAGAKU KOGYO KK.

XX WPI; 1995-242756/32.
 XX
 DR N-PSDB; AAT49535.

XX Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.
 XX
 PS Claim 3; Page 7-9; 38pp; English.

XX The sequences given in AA08986 represent the parvovirus structural
 CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.
 XX
 SQ Sequence 554 AA;

Query Match 100.0%; Score 118; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. NO. 3.2e-10; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLNMHTYFPNKGTOQYTDQIE 21
 |||||
 DB 386 GLNMHTYFPNKGTOQYTDQIE 406

RESULT 4
 AA23230
 ID AA23230 standard; Protein; 554 AA.
 XX
 AC AA23230;

DT 26-AUG-1999 (first entry)
 XX
 DE Erythrovirus V9 VP2 protein.

XX Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KW erythrovirus screening; typing; immunoassay; VP2 protein.

```

XX OS Erythrovirus.
XX PN FR2771751-A1.
XX PD 04-JUN-1999.
XX PF 03-DEC-1997; 97FR-0015197.
XX PR 03-DEC-1997; 97FR-0015197.
XX PA (ASST-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX PI Auguste V, Garbarg CA, Nguyen QT;
XX DR WPI: 1999-349543/30.
XX DR N-PSDB; ABX81586.
XX PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
    diagnosis of its infections
XX P Claim 19; Page 57-58; 80pp; French.
XX CC The present sequence represents an erythrovirus V9 protein.
XX CC Probes and primers derived from erythrovirus V9 polynucleotide
XX CC sequences (ABX81580) can be used for differential diagnosis of
XX CC erythrovirus (parvovirus) infections by a combination of
XX CC amplification and hybridisation assay. The probes can also be
XX CC used to assess susceptibility to erythrovirus infection and
XX CC for erythrovirus screening and typing. The antibodies can be
XX CC used in immunoassays for diagnosis of erythrovirus V9 infections.
XX SQ Sequence 554 AA;
Query Match 100.0%; Score 118; DB 20; Length 554;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLNMTTFPNKGTQGYTDQIE 21
DB 386 GLNMTTFPNKGTQGYTDQIE 406
RESULT 5
ABP57264
ID ABP57264 standard; Protein; 554 AA.
XX ABP57264;
XX 22-APR-2003 (first entry)
XX DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.
XX KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX OS Human parvovirus B19.
XX PN WO2003002753-A2.
XX PD 09-JAN-2003.
XX PF 28-JUN-2002; 2002WO-US20684.
XX PR 28-JUN-2001; 2001US-302077P.
XX PR 19-MAR-2002; 2002US-365956P.
XX PR 29-MAR-2002; 2002US-369224P.
XX PA (CHIR ) CHIRON CORP.
XX PI Pichuanes S, Shyamala V;
XX DR WPI: 2003-201510/19.
XX DR N-PSDB; ABZ59574.

```

```

XX PT Detecting a human parvovirus B19 infection in a biological sample to
XX PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
XX PT acid with a primer complementary to the 3'-terminal portion of the RNA
XX PT target sequence -
XX Example 4; Fig 7B; 148pp; English.
XX CC The present invention describes a method for detecting a human parvovirus
XX CC B19 infection in a biological sample. The method comprises reacting the
XX CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
XX CC consisting of a first primer containing a complexing sequence
XX CC sufficiently complementary to the 3'-terminal portion of the RNA target
XX CC sequence to complex with. Also described: (1) amplifying a target
XX CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
XX CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
XX CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
XX CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
XX CC consisting of a promoter region recognised by a DNA-dependent RNA
XX CC polymerase operably linked to a human parvovirus B19-specific complexing
XX CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
XX CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
XX CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
XX CC oligonucleotide primer of (4), and instructions for conducting the
XX CC diagnostic test. The method is useful for detecting parvovirus infection
XX CC in a biological sample, such as in blood products, to prevent
XX CC transmission of the virus through blood and plasma derivatives or by
XX CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
XX CC represent sequences used in the exemplification of the present invention.
XX SQ Sequence 554 AA;
Query Match 100.0%; Score 118; DB 24; Length 554;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLNMTTFPNKGTQGYTDQIE 21
DB 386 GLNMTTFPNKGTQGYTDQIE 406
RESULT 6
ABP57267
ID ABP57267 standard; Protein; 554 AA.
XX AC ABP57267;
XX DT 22-APR-2003 (first entry)
XX DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
XX KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX OS Human parvovirus B19.
XX PN WO2003002753-A2.
XX PD 09-JAN-2003.
XX PF 28-JUN-2002; 2002WO-US20684.
XX PR 28-JUN-2001; 2001US-302077P.
XX PR 19-MAR-2002; 2002US-365956P.
XX PR 29-MAR-2002; 2002US-369224P.
XX PA (CHIR ) CHIRON CORP.
XX PI Pichuanes S, Shyamala V;
XX DR WPI: 2003-201510/19.
XX DR N-PSDB; ABZ59577.
XX PT Detecting a human parvovirus B19 infection in a biological sample to

```

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -

PS Example 4; Fig 10B; 148pp; English.

CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.

SO Sequence 554 AA;

Query Match 100.0%; Score 118; DB 24; Length 554;
 Best Local Similarity 100.0%; Pred. No. 3.2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
 |||||
 DB 386 GLNMHTYFPNKGTOOYTDQIE 406

RESULT 7
 AAV71231
 ID AAV71231 standard; Protein; 756 AA.

AC AAV71231;

DT 08-SEP-2000 (first entry)

DE Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.

KW Adeno-associated virus; AAV2; human parvovirus B19; chimeric;
 KM recombinant parvoviral vector; cellular tropism; cap protein;
 KW capsid; gene delivery; gene therapy; VP1; VP2; VP3.

XX Chimeric - Adeno associated virus serotype 2.

OS Chimeric - Human parvovirus B19.

XX WO200028004-A1.

PN 18-MAY-2000.

PD 10-NOV-1999; 99WO-US26505.

XX 10-NOV-1998; 98US-0107840.

PR 10-MAR-1999; 99US-0123651.

XX (UNNC-) UNIV NORTH CAROLINA.

PA Rabinowitz JE, Samulski RJ, Xiao W;

XX WPI; 2000-376523/32.

DR N-PSDB; AAD00833.

PT Recombinant parvoviral vectors with altered packaging, tropisms and

PT immunogenic properties, useful in gene therapy protocols -
 XX Example 21; Page 142; 153pp; English.

CC The patent discloses modified parvovirus vectors with advantageous
 CC antigenic properties, packaging capabilities and cellular tropisms.
 CC These vectors can be used in standard recombinant DNA protocols e.g. gene
 CC therapy for delivering nucleic acids to cells.
 CC The present sequence is a capsid protein encoded by an
 CC adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric
 CC vector. This vector encodes AAV2 VP1 and VP2 capsid proteins
 CC and human parvovirus B19 VP2 protein. The chimeric vector was
 CC constructed by replacing the VP3 major cap protein of AAV2 with B19's
 CC VP2. Recombinant parvovirus comprising the chimeric capsid is useful
 CC for gene delivery.

SO Sequence 756 AA;

Query Match 100.0%; Score 118; DB 21; Length 756;
 Best Local Similarity 100.0%; Pred. No. 4.7e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
 |||||
 DB 588 GLNMHTYFPNKGTOOYTDQIE 608

RESULT 8
 AAM08986
 ID AAM08986 standard; Protein; 781 AA.

AC AAM08986;

DT 27-FEB-1997 (first entry)

DE Human parvovirus VP-1 protein.

KW Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KW erythroblastemia; abortion; universal fetal hydrops; liver disease;
 KW haemorrhagic fever; rheumatism; detection; IgG antibody.

XX Human parvovirus.

OS Key Location/Qualifiers

FT Misc-difference 19 /label= Gly, Val

FT Misc-difference 61 /label= Asn, Asp

FT Misc-difference 220 /label= His, Asn

PN JP07147986-A.

PD 13-JUN-1995.

PF 24-SEP-1992; 92JP-0281017.

PR 24-SEP-1992; 92JP-0281017.

XX (DENK-) DENKA SEIKEN KK.

PA (BLEED) DENKI KAGAKU KOGYO KK.

XX WPI; 1995-242756/32.

DR N-PSDB; AAT49535.

PT Human parvovirus gene coding for a polypeptide - useful for

XX developing vaccines against parvoviral diseases such as

PT erythroblastemia, haemorrhagic fever, etc.

XX Claim 2; Page 5-7; 38pp; English.

CC The sequences given in AAM08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythraemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

XX Sequence 781 AA;

Query Match 100.0%; Score 118; DB 16; Length 781;
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNNHTYFPNKGTOQYTDQIE 21
 |||||
 DB 613 GLNNHTYFPNKGTOQYTDQIE 633

RESULT 9

AA23227 standard; Protein; 781 AA.

AC AA23227;

DT 26-AUG-1999 (first entry)

DE Erythrovirus V9 VP1 protein.

XX Erythrovirus V9; differential diagnosis; parvovirus; infection;

KM erythrovirus screening; typing; immunoassay; VP1 protein.

XX Erythrovirus.

PN FR2771751-A1.

PD 04-JUN-1999.

PF 03-DEC-1997; 97FR-0015197.

PR 03-DEC-1997; 97FR-0015197.

PA (ASST-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

PI Auguste V, Garbarg CA, Nguyen QT;

XX WPI: 1999-349543/30.

DR N-PSDB; AAX81583.

XX Erythrovirus V9 and its nucleic acid sequences - can be used in the
 diagnosis of its infections

PS Claim 19; Page 50-52; 80pp; French.

XX The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.

XX Sequence 781 AA;

Query Match 100.0%; Score 118; DB 20; Length 781;
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNNHTYFPNKGTOQYTDQIE 21
 |||||
 DB 613 GLNNHTYFPNKGTOQYTDQIE 633

RESULT 10

ABP57263 standard; Protein; 781 AA.

XX ABP57263;

DT 22-APR-2003 (first entry)

DE Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

OS Human parvovirus B19.

PN WO2003002753-A2.

PD 09-JAN-2003.

PF 28-JUN-2002; 2002WO-US20684.

PR 28-JUN-2001; 2001US-302077P.

PR 19-MAR-2002; 2002US-365956P.

PR 29-MAR-2002; 2002US-369224P.

PA (CHIR) CHIRON CORP.

PI Pichantes S, Shyamala V;

DR WPI: 2003-201510/19.

DR N-PSDB; ABZ59573.

PT Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -

XX Example 4; Fig 6B; 148pp; English.

PS The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABZ57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 118; DB 24; Length 781;
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNNHTYFPNKGTOQYTDQIE 21
 |||||
 DB 613 GLNNHTYFPNKGTOQYTDQIE 633

RESULT 11
 ABP57266

```

ID  ABP57266 standard; Protein; 781 AA.
XX
XX  ABP57266;
AC
XX  22-APR-2003 (first entry)
XX
XX  Human parvovirus B19 clone B6-VP1 amino acid sequence SEQ ID NO:33.
XX
XX  Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
OS
XX  Human parvovirus B19.
XX
XX  WO2003002753-A2.
XX
XX  09-JAN-2003.
XX
XX  28-JUN-2002; 2002WO-US20684.
XX
XX  28-JUN-2001; 2001US-302077P.
XX  19-MAR-2002; 2002US-365956P.
XX  29-MAR-2002; 2002US-369224P.
XX
XX  (CHIR ) CHIRON CORP.
XX
XX  Pichuanee S, Shyamala V;
XX
XX  WPI, 2003-201510/19.
XX  N-PSDB; ABZ59576.
XX
XX  Detecting a human parvovirus B19 infection in a biological sample to
XX  prevent viral transmission, comprises reacting a parvovirus B19 nucleic
XX  acid with a primer complementary to the 3'-terminal portion of the RNA
XX  target sequence -
XX
XX  Example 4; Fig 9B; 148pp; English.
XX
XX  The present invention describes a method for detecting a human parvovirus
XX  B19 infection in a biological sample. The method comprises reacting the
XX  isolated parvovirus B19 nucleic acid with a first oligonucleotide
XX  consisting of a first primer containing a complexing sequence
XX  sufficiently complementary to the 3'-terminal portion of the RNA target
XX  parvovirus B19 nucleotide sequence; (1) amplifying a target
XX  parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
XX  of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
XX  ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
XX  sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
XX  consisting of a promoter region recognised by a DNA-dependent RNA
XX  polymerase operably linked to a human parvovirus B19-specific complexing
XX  sequence of 10-75 nucleotides; (5) an oligonucleotide probe complexing a
XX  parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
XX  to an acridinium ester label; and (6) a diagnostic test kit comprising an
XX  oligonucleotide primer of (4), and instructions for conducting the
XX  diagnostic test. The method is useful for detecting parvovirus infection
XX  in a biological sample, such as in blood products, to prevent
XX  transmission of the virus through blood and plasma derivatives or by
XX  close personal contact. ABZ59559 to ABZ59634 and ABP57262 to ABP57267
XX  represent sequences used in the exemplification of the present invention.
XX
XX  Sequence 781 AA;
XX
XX  Query Match 100.0%; Score 118; DB 24; Length 781;
XX  Best Local Similarity 100.0%; Pred. No. 4.9e-10;
XX  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  1 GIANHTYFPNKGTQOYTDOIE 21
XX  |||||
XX  613 GIANHTYFPNKGTQOYTDOIE 633
XX
XX  RESULT 12
XX  ID ABP57712
XX  ABP57712 standard; Protein; 93 AA.
XX

```

AC	ABP57712;
XX	
DT	22-JAN-2003 (first entry)
XX	
DE	Zinc finger protein FPM31510.23.
XX	
KM	Zinc finger protein; FPM31510.23; embryo development; tumour; diabetes.
XX	
OS	Unidentified.
XX	
PN	CN1351017-A.
XX	
PD	29-MAY-2002.
XX	
PF	26-OCT-2000; 2000CN-0125780.
XX	
PR	26-OCT-2000; 2000CN-0125780.
XX	
PA	(BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX	
PI	Mao Y, Xie Y;
XX	
DR	WPI; 2002-644398/70.
DR	N-PESDB; ABV75581.
XX	
PT	Poly peptide-zinc finger protein FPM 31510.23 and polynucleotide for coding it -
XX	
PS	Claim 1; Page 27 (Disclosure); 33pp; Chinese.
CC	The invention relates to a novel polypeptide, zinc finger protein FPM31510.23, and the polynucleotide encoding it. The polypeptide is useful for treating diseases such as embryo development deformity, CC tumour, and diabetes. The antagonist of the polypeptide and its medical action, and the application of the polynucleotide are also disclosed. The present sequence represents the zinc finger protein FPM31510.23 of the invention.
CC	
CC	
CC	
SQ	Sequence 93 AA;
Query Match	39.8%; Score 47; DB 23; Length 93;
Best Local Similarity	50.0%; Pred. No. 8.7;
Matches	7; Conservative 3; Mismatches 4; Indels 0; Gaps 0.
Oy	1 GLNMHTYFPNKGTQ 14 : : : : : :
Db	79 GINLHLVPESEFTE 92
RESULT 13	
ABB48499	
ID	ABB48499 standard; Protein; 305 AA.
XX	
NC	ABB48499;
DT	05-FEB-2002 (first entry)
XX	
DE	Listeria monocytogenes protein #1203.
XX	
KM	Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation; vitamin B12; bacterial infection; disease.
XX	
OS	Listeria monocytogenes.
XX	
PN	MO200177335-A2.
XX	
PD	18-OCT-2001.
XX	
PF	11-APR-2001; 2001WO-FR01118.
XX	
DR	11-APR-2000; 2000FR-0004629.
XX	
PA	(INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve B, Rusniok C, Fshli H, Dehoux P;
 PI Duesuguet O, Chetoui F, Medjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baghero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Mahueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX WPI; 2002-010914/01.
 XX
 PT Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 PT and prevention of *Listeria* and related bacterial infections, and
 PT related polypeptides -
 XX
 PS Claim 6; SEQ ID No 1204; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of *Listeria*
 CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC *monocytogenes* and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC *monocytogenes* and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 305 AA;
 XX
 Query Match 39.8%; Score 47; DB 23; Length 305;
 Best Local Similarity 37.5%; Pred. No. 36;
 Matches 9; Conservative 5; Mismatches 6; Indels 4; Gaps 1;
 QY 1 GLNHTYFPN---KGTQGYTDQI 20
 Db 93 GFNHTYFPDWMQESIEQITRL 116
 XX
 UT 14
 ID ABL13993 standard; Protein; 904 AA.
 AC ABL13993;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #13984.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.

XX
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS78180.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 44352; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAG00010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 904 AA;
 XX
 Query Match 39.8%; Score 47; DB 22; Length 904;
 Best Local Similarity 41.2%; Pred. No. 1.4e+02;
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 4 MHTYFPNKGTYQYTDQI 20
 Db 166 LMTYAPNKGASRPTKQV 182
 XX
 RESULT 15
 ID AAW98428 standard; Protein; 225 AA.
 AC AAW98428;
 XX
 DT 31-MAR-1999 (first entry)
 XX
 DE *H. pylori* GHP0 409 protein.
 XX
 KW GHP0 protein; *Helicobacter* infection; gastroduodenal disease; gastritis;
 KW peptic ulcer disease.
 XX
 OS *Helicobacter pylori*.
 XX
 PN WO9843478-A1.
 XX
 PD 08-OCT-1998.
 XX
 PF 01-APR-1998; 98WO-US06371.
 XX
 PR 29-JUL-1997; 97US-0902615.
 PR 01-APR-1997; 97US-0833457.
 PR 24-JUN-1997; 97US-0881227.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 767 AA;

Query Match 38.1%; Score 45; DB 22; Length 767;
 Best Local Similarity 53.8%; Pred. No. 2.3e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 8 FPNKGTQOYTDQI 20
 ||::|||
 511 FPNRGTSTYADQM 523

RESULT 18

AAV11067
 ID AAV11067 standard; Protein; 814 AA.

AAV11067;

DT 08-JUN-1999 (first entry)

DE H. pylori ORF 13ae10610_6522827_c3_37 outer membrane protein.

KM Vaccine; probe; diagnostic; ORF; cell envelope protein;

KW secreted protein; cytoplasmic protein; cellular protein.

XX Helicobacter pylori.

XX WO9824475-A1.

XX 11-JUN-1998.

XX 05-DEC-1997; 97WO-US22104.

XX 14-JUL-1997; 97US-0891928.

XX 05-DEC-1996; 96US-0759625.

XX 25-MAR-1997; 97US-0823745.

XX (ASTR) ASTRA AB.

PI Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;

XX WPI; 1998-333051/29.

XX N-PSDB; AAX30596.

PT New isolated Helicobacter pylori nucleic acids - used to develop
 products for the diagnosis, prevention and treatment of infection by
 H. pylori and other Helicobacter species

XX Claims 37, 41; Page 230-232; 339pp; English.

CC Recombinant or substantially pure preparations of H. pylori polypeptides
 CC are disclosed, together with the nucleic acids encoding them. In all,
 CC 97 ORFs are shown. The proteins are variously cell envelope proteins,
 CC cytoplasmic proteins, secreted proteins or other cellular proteins.
 CC Vaccines containing the nucleic acids or proteins are claimed, as are
 CC probes containing at least 8 nucleotides from the nucleic acid
 CC sequences. The vaccines are useful for treating or reducing the risk of
 CC H. pylori infections, and the probes can be used diagnostically for
 CC detecting the presence of Helicobacter in a sample. The products are
 CC also of use in screening for compounds having the ability to interfere
 CC with the H. pylori life cycle or to inhibit H. pylori infection.

XX Sequence 814 AA;

Query Match 38.1%; Score 45; DB 19; Length 814;

Best Local Similarity 47.4%; Pred. No. 2.5e+02;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 GLNNHTYFPNKGTOOYTDQ 19

Db ||::|||
 762 GLSLNAVFNNVFNQOYTDQ 780

RESULT 19

AAV17162
 ID AAV17162 standard; Protein; 814 AA.

AAV17162;

DT 03-AUG-1999 (first entry)

DE H. pylori outer membrane polypeptide.

KM Outer membrane polypeptide; OMP; vaccine; H. pylori infection; humoral;

KW cellular immune response.

XX Helicobacter pylori.

XX WO9921959-A2.

XX 06-MAY-1999.

XX 28-OCT-1998; 98WO-US22883.

XX 17-DEC-1997; 97US-0993001.

XX 28-OCT-1997; 97US-0959131.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Alm RA, Ellis RW, Guild BC, Noonan BM, Smith D;

XX WPI; 1999-326698/27.

XX N-PSDB; AAX75781.

PT Cellular vaccine against Helicobacter pylori

CC Claim 7; Page 198-202; 352pp; English.
 CC The invention relates to a vaccine for preventing or treating infections
 CC by Helicobacter pylori. The vaccine contains at least one isolated
 CC H. pylori polypeptide, or its fragments, in a carrier, where the
 CC carrier is a Salmonella, Vibrio cholerae or Shigella vector containing a
 CC nucleic acid encoding the H. pylori polypeptide. The vaccines induce
 CC humoral and cellular immune responses. The vaccines are used to treat or
 CC prevent infections by H. pylori. Sequences AAX75779 to AAX75837 represent
 CC nucleic acid sequences encoding H. pylori outer membrane polypeptides
 CC (OMPs) AAV17160 to AAV17218.

XX Sequence 814 AA;

Query Match 38.1%; Score 45; DB 20; Length 814;
 Best Local Similarity 47.4%; Pred. No. 2.5e+02;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 GLNNHTYFPNKGTOOYTDQ 19
 ||::|||
 Db 762 GLSLNAVFNNVFNQOYTDQ 780

RESULT 20

AAV11066
 ID AAV11066 standard; Protein; 815 AA.

AAV11066;

DT 08-JUN-1999 (first entry)

DE H. pylori ORF 13ae10610_156411_c3_33 outer membrane protein.

KM Vaccine; probe; diagnostic; ORF; cell envelope protein;

XX secreted protein; cytoplasmic protein; cellular protein.

DR WPI: 2000-602362/57.
 DR N-PSDB; AAC75507.
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 1590-1591; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteoplastic; anticonvulsant; antiallergic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antitumor; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease, to enhance
 CC coagulation, to inhibit thrombosis, and as a contraceptive.
 CC
 SQ Sequence 142 AA;
 XX
 QY 3 NMHTYFP---NKGTQYTDQIE 21
 Db 7 NLHPFAPAEBSAGYQQLTDELE 28
 XX
 RESULT 23
 ABP07047
 ID ABP07047 standard; Protein; 142 AA.
 XX
 AC ABP07047;
 XX
 24-JUN-2002 (first entry)
 DE Human ORFX protein sequence SEQ ID NO:14076.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 KW
 OS Homo sapiens.
 OS
 XX
 PN WO200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US10836.
 XX
 PR 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 XX
 PA (CURA-) CURAGEN CORP.

XX
 PI Shimketa RA, Leach MD;
 XX
 DR WPI: 2002-106308/14.
 DR N-PSDB; ABN22799.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 PS Disclosure; SEQ ID 14076; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 142 AA;
 XX
 QY 3 NMHTYFP---NKGTQYTDQIE 21
 Db 7 NLHPFAPAEBSAGYQQLTDELE 28
 XX
 RESULT 24
 ABB61725
 ID ABB61725 standard; Protein; 406 AA.
 XX
 AC ABB61725;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 11967.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW
 OS Drosophila melanogaster.
 OS
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX

PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI, 2001-656660/75.
 DR N-PSDB; ABL05828.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Disclosure; SEQ ID NO 11967; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 406 AA;
 OY 4 MHTYPPN-----KGTQGYTDQIE 21
 Db 198 LHTYPPNIFDIKYLKMSCKNLKGLQEVADQIE 230
 RESULT 25
 AAM48920
 ID AAM48920 standard; Protein; 204 AA.
 XX
 AC AAM48920;
 XX
 DT 05-APR-2002 (first entry)
 XX
 DE Human L1 retroposon 22.
 XX
 HM Human; L1 retroposon 22; cancer; Duchenne muscular dystrophy;
 XX haemopathy; developmental disorder; HIV infection; immunological disease;
 XX inflammation; cytostatic; haemostatic; virucide; immunomodulatory;
 XX antiinflammatory; gene therapy.
 OS Homo sapiens.
 XX
 PN WO200194404-A1.
 XX
 PD 13-DEC-2001:
 XX
 PF 14-MAY-2001; 2001WO-CN00761.
 XX
 XX 16-MAY-2000; 2000CN-0115705.
 XX
 PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI; 2002-090140/12.
 DR N-PSDB; AAK98483.
 XX
 PT Human L1 retroposon 22 and encoding polynucleotide, used in diagnosis
 PT and treatment of malignant tumors, hemopathy, human immunodeficiency
 PT virus infection, immunological diseases and inflammation -
 PT
 Claim 1; Page 29; 36bp; Chinese.

Query Match	Similarity	Score	DB	Length
Best Local	91.4%	37.38	47.48	204
Matches	9	Conservative	3	Mismatches
			5	Indels
			2	Gaps
			1	

2 LNMHTYFPNKGSTQOYTDOI 20
 104 LNMHT--PNTGASTFTRKQL 120

RESULT 26
 ABP73753
 ID ABP73753 standard; Protein; 1017 AA.
 AC ABP73753;
 DT 30-JAN-2003 (first entry)
 DE Candida albicans essential protein SEQ ID NO 7590.
 XX
 KW Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
 KW signal transduction; DNA replication; cell division; growth;
 XX proliferation; Candida albicans; fungicide; antifungal.
 OS Candida albicans.
 XX
 PN WO200253728-A2.
 XX
 PD 11-JUL-2002.
 XX
 PE 26-DEC-2001; 2001WO-US949486.
 XX
 PR 29-DEC-2000; 2000US-259128P.
 PR 20-FEB-2001; 2001US-0782024.
 PR 22-AUG-2001; 2001US-314050P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
 PI WPI: 2002-566694/60.
 DR N-PSDB; ABZ32303.
 XX
 CC Constructing strains for identifying gene products as effective targets
 CC for therapeutic intervention, by inactivating in the strain one allele
 CC of a gene and placing other allele of the gene under conditional
 CC expression -
 CC
 CC Claim 44; SEQ ID NO 7590, 1677p + Sequence Listing; English.
 CC
 CC The invention relates to constructing (M1) a strain of diploid fungal
 CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination, of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells in which both alleles of a gene are modified. The diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the

CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthetic, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of C. albicans cells and for
 CC treating infection by C. albicans. The present sequence is that of an
 CC essential Candida albicans protein used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.

XX
 SQ Sequence 1017 AA;

Query Match 37.3%; Score 44; DB 23; Length 1017;
 Best Local Similarity 50.0%; Pred. No. 4.7e+02;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GLNMHTYFPNKGTOQY 16
 |||||
 591 GLNSELVYFONKMSKY 606

RESULT 27

AB93919
 ID ABB93919 standard; Protein; 616 AA.

AC ABB93919;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 3130.

KM Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 26-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -

PS Claim 5; SEQ ID NO 3130; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
 CC (ABB9790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

XX Sequence 616 AA;

Query Match 36.9%; Score 43.5; DB 23; Length 616;
 Best Local Similarity 42.9%; Pred. No. 3.1e+02;
 Matches 9; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

Qy 2 LNMHTYFPNKGTOQYTDQ 19
 |||||
 Db 292 LNMHTYFPNKGTOQYTDQ 312

RESULT 28

ID AAW21827 standard; Protein; 119 AA.

AC AAW21827;

DT 26-OCT-1997 (first entry)

DE Rat RT1.Balplal class II MHC alpha chain.

KM Immunosuppression; major histocompatibility complex class II;
 KM MHC; mixed lymphocyte reaction; allorecognition; cytotoxic T cell;
 KM alloimmunity; autoimmune disease; organ transplant;
 KM multiple sclerosis; rheumatoid arthritis; RT1.Balplal.

OS Rattus sp.

FH Key Location/Qualifiers

FT Protein 25..119 /note= "alpha chain amino acids 1-95"

FT Region 87..103 /note= "alpha chain aa62-79 conserved region
 (Claim 12)"

PN WO9710711-A1.

PD 27-MAR-1997.

PF 23-SEP-1996; 96NO-US15662.

PR 21-SEP-1995; 95US-0004117.

PA (AUTO-) AUTOIMMUNE INC.

PI Carpenter CB, Murphy BT, Sayegh MH;

XX WPI; 1997-202534/18.

PT Suppression of immune responses with major histocompatibility
 PT complex class II peptide(s) - useful in alloimmunity, e.g. organ
 PT transplantation, and autoimmunity, e.g. in multiple sclerosis or
 PT rheumatoid arthritis

PS Claim 10; Fig 2A-B; 55pp; English.

XX The sequence includes the first 95 amino acid residues of the
 CC alpha chain of the class II MHC from rat RT1.Balplal. It contains
 CC a highly conserved region that is maintained across alleles and
 CC species (see also AAW21828 and AAW21829). A novel method of suppressing
 CC an immune response comprises administering a class II MHC alpha
 CC chain, or a fragment (see also AAW21830-32), that can suppress at
 CC least one of the following: a mixed lymphocyte reaction or other
 CC T-cell allorecognition reaction; generation of cytotoxic T-cells
 CC recognising an alloantigen; lymphocyte proliferation against tissue
 CC antigen; and stimulatory cytokine production by lymphocytes. Immune
 CC responses that can be abated or suppressed include alloimmunity
 CC (e.g. in organ transplantation) and autoimmunity (e.g. in multiple
 CC sclerosis or rheumatoid arthritis). Immune responses can be down-
 CC regulated specifically without the adverse effects of conventional
 CC treatments.

XX Sequence 119 AA;

Query Match 36.4%; Score 43; DB 18; Length 119;
 Best Local Similarity 33.3%; Pred. No. 51;
 Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

Qy 1 GLNMHTYFPNKGTOQYTDQIE 21

Db 36 GINMYQYESRG--QFTHERD 54

RESULT 29

AA06804
ID AA06804 standard; Protein; 180 AA.

AC AA06804;

DT 23-JUN-1999 (first entry)

DE Peptide Seq ID No: 24.

Major histocompatibility complex; MHC; antigen binding cleft; allergy; antigen-specific T-cell; transplant rejection; autoimmune disease; multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus; insulin-dependent diabetes mellitus; immune response; cancer.

Rattus sp.

MO9914236-A1.

25-MAR-1999.

15-SEP-1998; 98WO-US18244.

10-OCT-1997; 97US-0064555.

16-SEP-1997; 97US-0064552.

(BURR/) BURROWS G G.

(VAND/) VANDENBARK A A.

Burrows GG, Vandenbark AA;

WPI; 1999-229498/19.

Recombinant 2-domain major histocompatibility complex molecules used to manipulate antigen specific T-cells

Example 1; Page 69-70; 73pp; English.

The invention relates to recombinant polypeptides comprising only those domains of major histocompatibility complex (MHC) molecules that define the antigen binding cleft. These polypeptides can be used to mimic the function of mammalian MHC. The recombinant polypeptide comprises covalently linked first and second domains, where the first domain is a mammalian MHC class II alpha domain, or where the first domain is a mammalian MHC class I alpha domain and the second domain is a mammalian MHC class I alpha2 domain, and wherein the polypeptide is not a whole MHC class I alpha chain. The amino terminus of the second domain is covalently linked to the carboxy terminus of the first domain in both cases, optionally with a peptide linker sequence. The recombinant MHC polypeptides are useful for a wide range of in vitro and in vivo applications, and may be used in place of either intact, purified MHC molecules or antigen presenting cells that express MHC molecules. The polypeptides can be used for the detection, quantification and purification of antigen-specific T-cells in biological samples. They can also be used to activate or inactivate T-cells and to induce energy in T-cells. The polypeptides may be used for the amelioration of conditions mediated by antigen-specific T-cells, e.g. allergies, transplant rejection and autoimmune diseases including multiple sclerosis, rheumatoid arthritis, systemic lupus erythematosus and insulin-dependent diabetes mellitus. The polypeptides may also be used to boost immune responses in certain conditions such as cancer and infectious diseases. If a toxic molecule is attached to the polypeptides, then they can also be used to kill specific T-cells.

Sequence 180 AA;

Query Match 36.4%; Score 43; DB 20; Length 180;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 1 GINMYTYPNKGTOQYTDQIE 21
Db 105 GINMYQYESRG--QFTHERD 123

RESULT 30

AAE33454
ID AAE33454 standard; Protein; 180 AA.

AC AAE33454;

DT 02-APR-2003 (first entry)

DE Rat RT1B beta2alpha protein #2.

Major histocompatibility complex; MHC; therapy; immune-mediated disorder; insulin-dependent diabetes mellitus; inflammatory bowel disease; uveitis; polyarteritis; multiple sclerosis; myasthenia gravis; antiinflammatory; rheumatoid arthritis; neuroprotective; thyroiditis; ophthalmological; chronic beryllium disease; rat.

Rattus sp.

Key Location/Qualifiers

Domain 1..97 /note="Beta1 domain"

Domain 98..180 /note="Alpha1 domain"

MO200287613-A1.

07-NOV-2002.

01-MAY-2002; 2002WO-US13873.

01-MAY-2001; 2001US-0847172.

(UVOR-) UNIV OREGON HEALTH SCI.

Burrows GG, Vandenbark AA;

WPI; 2003-103435/09.

New purified MHC class II polypeptide with covalently linked beta1 and alpha1 domains, useful for treating or preventing an immune-mediated disorder such as rheumatoid arthritis, diabetes mellitus and inflammatory bowel disease.

Disclosure; Page 125-126; 133pp; English.

The invention relates to major histocompatibility complex (MHC) class II polypeptide comprising covalently linked beta1 and alpha1 domains. The methods and compositions of the present invention are useful for treating or preventing an immune-mediated disorder such as rheumatoid arthritis, chronic beryllium disease, thyroiditis, insulin-dependent diabetes mellitus, inflammatory bowel disease, uveitis, polyarteritis, multiple sclerosis and myasthenia gravis. The present sequence is rat RT1B beta2alpha protein used in the invention.

Sequence 180 AA;

Query Match 36.4%; Score 43; DB 24; Length 180;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 1 GINMYTYPNKGTOQYTDQIE 21
Db 105 GINMYQYESRG--QFTHERD 123

RESULT 31
AA06797

ID AAY06797 standard; Protein; 185 AA.
 AC AAY06797;
 DT 23-JUN-1999 (first entry)
 XX
 DE Betalaphal/MBP-72-89 antigen/linker insert peptide sequence.
 KW Major histocompatibility complex; MHC; antigen binding cleft; allergy;
 KW antigen-specific T-cell; transplant rejection; autoimmune disease;
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;
 KW insulin-dependent diabetes mellitus; immune response; cancer.
 OS Synthetic.
 XX
 XX MO9914236-A1.
 XX
 XX 25-MAR-1999.
 XX
 XX 15-SEP-1998; 98WO-US18244.
 XX
 XX 10-OCT-1997; 97US-0064555.
 XX 16-SEP-1997; 97US-0064552.
 XX
 XX (BURR/) BURROWS G G.
 XX (VAND/) VANDENBARK A A.
 XX
 XX Burrows GG, Vandenbark AA;
 XX
 XX MPI: 1999-229498/19.
 XX DR N-PSDB: AAX32524.
 XX
 PT Recombinant 2-domain major histocompatibility complex molecules used
 PT to manipulate antigen specific T-cells
 PS
 PS Example 1; Fig 1B; 73pp: English.
 XX
 CC The invention relates to recombinant polypeptides comprising only those
 CC domains of major histocompatibility complex (MHC) molecules that define
 CC the antigen binding cleft. These polypeptides can be used to mimic the
 CC function of mammalian MHC. The recombinant polypeptide comprises
 CC covalently linked first and second domains, where the first domain is a
 CC mammalian MHC class II beta domain and the second domain is a mammalian
 CC Class I alpha domain, or where the first domain is a mammalian MHC
 CC Class I alpha domain and the second domain is a mammalian MHC class I
 CC alpha2 domain, and wherein the polypeptide is not a whole MHC class I
 CC alpha chain. The amino terminus of the second domain is covalently linked
 CC to the carboxy terminus of the first domain in both cases, optionally
 CC with a peptide linker sequence. The recombinant MHC polypeptides are
 CC useful for a wide range of in vitro and in vivo applications, and may be
 CC used in place of either intact, purified MHC molecules or antigen
 CC presenting cells that express MHC molecules. The polypeptides can be used
 CC for the detection, quantification and purification of antigen-specific
 CC T-cells in biological samples. They can also be used to activate or
 CC inactivate T-cells and to induce energy in T-cells. The polypeptides may
 CC be used for the amelioration of conditions mediated by antigen-specific
 CC T-cells, e.g. allergies, transplant rejection and autoimmune diseases
 CC including multiple sclerosis, rheumatoid arthritis, systemic lupus
 CC erythematosus and insulin-dependent diabetes mellitus. The polypeptides
 CC may also be used to boost immune responses in certain conditions such as
 CC cancer and infectious diseases. If a toxic molecule is attached to the
 CC polypeptides, then they can also be used to kill specific T-cells. The
 CC present sequence represents an antigen/linker insert sequence suitable
 CC for insertion into a single chain betalaphal expression cassette of the
 CC invention.
 CC
 XX
 XX Sequence 185 AA;
 XX
 XX Query Match 36.4%; Score 43; DB 20; Length 185;
 XX Best Local Similarity 33.3%; Pred. No. 87;
 XX Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

```

Db          110 GINMYQYSESG--QTFHEFD 128

RESULT 32
AAE33447
ID AAE33447 standard; Protein; 185 AA.
XX AC
XX AAE33447;
DT 02-APR-2003 (first entry)
DE Rat RT1B beta1alpha1 protein #1.
XX Major histocompatibility complex; MHC; therapy; immune-mediated disorder;
XX insulin-dependent diabetes mellitus; inflammatory bowel disease; uveitis;
XX polyarteritis; multiple sclerosis; myasthenia gravis; arthralgic;
XX rheumatoid arthritis; neuroprotective; thyroiditis; ophthalmological;
XX chronic beryllium disease; rat.
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Domain 1..95
FT /note= "Beta1 domain"
FT Domain 96..185
FT /note= "Alpha1 domain"
XX
XX W0200287613-A1.
XX
PD 07-NOV-2002.
XX
PF 01-MAY-2002; 2002MO-US13873.
XX
PR 01-MAY-2001; 2001US-0847172.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Burrows GG, Vandenbark AA;
XX
DR WPI; 2003-103435/09.
XX
N-PSDB; AAD51190.
XX
PT New purified MHC class II polypeptide with covalently linked beta1 and
PT alpha domains, useful for treating or preventing an immune-mediated
PT disorder such as rheumatoid arthritis, diabetes mellitus and.
PT inflammatory bowel disease -
XX
PS Example 1; Page 117-118; 133pp; English.
XX
XX The invention relates to major histocompatibility complex (MHC) Class
XX II polypeptide comprising covalently linked beta1 and alpha domains.
XX The methods and compositions of the present invention are useful for
XX treating or preventing an immune-mediated disorder such as rheumatoid
XX arthritis, chronic beryllium disease, thyroiditis, insulin-dependent
XX diabetes mellitus, inflammatory bowel disease, uveitis, polyarteritis,
XX multiple sclerosis and myasthenia gravis. The present sequence is rat
XX RT1B beta1alpha1 protein. This sequence is used in the exemplification
XX of the invention.
XX
SQ Sequence 185 AA;

Query Match 36.4%; Score 43; DB 24; Length 185;
Best Local Similarity 33.3%; Pred. No. 87;
Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

CY 1 GUNMHITYFENKGTQOYTDOIIE 21
|::||:|::|::|:|:|
|::||:|::|::|:|:|
Db          110 GINMYQYSESG--QTFHEFD 128
ABG02830
```

ID ABG02830 standard; Protein; 204 AA.
 XX
 AC ABG02830;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #2821.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PP 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX
 PA (HSE-) HSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 XX N-PSDB; AAS67017.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 33189; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence. 204 AA;
 XX
 QY Query Match. 36.4%; Score 43; DB 22; Length 204;
 DB Best Local Similarity 38.5%; Pred. No. 98;
 Matches 10; Conservative 4; Mismatches 6; Indels 6; Gaps 1;
 XX
 1 GLNHTYFPNKG-----TQGYTDQ 20
 61 GSTWQTFEPTGELTCPEPTQGTDEV 86
 XX
 RESULT 34
 AAR72588
 ID AAR72588 standard; Protein; 297 AA.
 XX

AC AAR72588;
 XX
 DT 25-MAR-2003 (updated)
 DT 23-NOV-1995 (first entry)
 XX
 DE Carponin protein.
 XX
 KW Carponin; arteriosclerosis; ischaemia; heart; PTCA.
 XX
 OS Homo sapiens.
 XX
 PN WO9509010-A1.
 XX
 PD 06-APR-1995.
 XX
 PF 28-FEB-1994; 94WO-JP00320.
 XX
 PR 29-SEP-1993; 93JP-0243381.
 XX
 PA (KIRI) KIRIN BREWERY KK.
 PA (OSAP) OSAKA PREFECTURE.
 PA (KIRI) KIRIN BEER KK.
 PA (OSAP) OSAKA PREFECTURAL GOVERNMENT.
 XX
 PI Shibata N, Takahashi K;
 XX
 DR WPI; 1995-147255/19.
 DR N-PSDB; AAO86719.
 XX
 PT Sequence encoding carponin gene and expression vector comprising
 PT it - useful for the treatment of arteriosclerosis and ischaemic
 PT heart disorders
 XX
 PS Claim 2; Page 19-20; 35bp; Japanese.
 XX
 CC Carponin inhibits the thickening of the endangium. Carponin may
 CC be used in the treatment of other ischaemic heart disorders such
 CC as angina pectoris and myocardial infarction. Carponin may also
 CC be used after percutaneous transluminal coronary angioplasty (PTCA)
 CC procedures to prevent restenosis.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX
 SQ Sequence 297 AA;
 XX
 QY Query Match. 36.4%; Score 43; DB 16; Length 297;
 DB Best Local Similarity 47.4%; Pred. No. 1.5e+02;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 XX
 1 GLNHTYFPNKGTTQGYTDQ 19
 183 GTRHLYDPKLGTDQPLDQ 201
 XX
 RESULT 35
 AAR94888
 ID AAR94888 standard; Protein; 297 AA.
 XX
 AC AAR94888;
 XX
 DT 25-SEP-1996 (first entry)
 DT Carponin.
 DE Carponin.
 XX
 KW carponin; anticancer agent; metastasis inhibitor.
 XX
 OS Homo sapiens.
 XX
 PN JF08073380-A.
 XX
 PD 19-MAR-1996.
 XX
 PF 02-SEP-1994; 94UP-0232498.
 XX

```

PR      02-SEP-1994;    94JP-0232498.
XX
PA      (KIRI ) KIRIN BREWERY KK.
PA      (OSAP ) OSAKA PREFECTURE.
XX
DR      MPI; 1996-205463/21.
DR      N-PSDB; AAT16663.
XX
PT      Anticancer agent comprising the carponin gene as active ingredient -
PT      lowers tumour-forming activity of cancer cells and reduces
PT      metastatic activity
XX
PS      Claim 4; Page 9-10; 14pp; Japanese.
XX
CC      The present sequence is that of carponin which is useful as an
CC      anticancer agent, esp. as an inhibitor of metastasis. The agent
CC      lowers the tumour-forming activity of the cancer cells.
XX
SQ      Sequence . 297 AA;
●
CY      Try Match          36.4%; Score 43; DB 17; Length 297;
        Local Similarity   47.4%; Pred. NO. 1.5e+02;
Matches  9; Conservative  0; Mismatches  10; Indels  0; Gaps  0,
        1 GLNMHTYFPNKGTOQYTQD 19
Db       183 GTRRRLYPKLGTDPDLQ 201

RESULT 36
ABB65502
ID      ABB65502 standard; Protein; 493 AA.
AC      ABB65502;
XX
DT      26-MAR-2002 (first entry)
DE      Drosophila melanogaster polypeptide SEQ ID NO 23298.
XX
KW      Drosophila; developmental biology; cell signalling; insecticide;
KW      pharmaceutical.
XX
OS      Drosophila melanogaster.
XX
WO200171042-A2.
XX
PN      WO200171042-A2.
PD      27-SEP-2001.

23-MAR-2001; 2001WO-US09231.
PR      23-MAR-2000; 2000US-191637P.
PR      11-JUL-2000; 2000US-0614150.
XX
PA      (PEKE ) PE CORP NY.
XX
PI      Venter JC, Adams M, Li PWD, Myers EW;
XX
WI      MPI; 2001-656860/75.
DR      N-PSDB; ABL09605.
XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
PT      interactions -
XX
PS      Disclosure: SEQ ID NO 23298; 21pp + Sequence listing; English.
XX
CC      The invention relates to an isolated nucleic acid detection reagent
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
CC      useful in developmental biology and in elucidating cell signalling and
CC      cell-cell interactions in higher eukaryotes for the development of
CC      insecticides, therapeutics and pharmaceutical drugs. The invention
CC      discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC      sequences (ABL01840-ABL16175) and the encoded proteins

```

CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC	
CC	
XX	
SQ	Sequence 493 AA;
OY	2 LNMHTYPNKGTOOYT--DOJE 21 :: : : Best Local Similarity 36.4%; Score 43; DB 22; Length 493; Matches 8; Conservative 5; Mismatches 7; Indels 2; Gaps 1
Dd	191 VSFHYFPNQMFETVPYEAB 212
RESULT 37	
ABB77613	
ID	ABB77613 standard; Protein; 726 AA.
AC	ABB77613;
XX	
DT	01-JUL-2002 (first entry)
XX	
DE	AMEPV NTPase (AMV087).
KW	AMEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
KX	genetic deficiency disorder; NTPase.
OS	Amaacta moorei entomopoxvirus.
XX	
EN	WO200212526-A2.
XX	
PD	14-FEB-2002.
XX	
PF	10-AUG-2001; 2001WO-US25287.
XX	
PR	10-AUG-2000; 2000US-224479P.
XX	
PR	14-SEP-2000; 2000US-0662254.
PA	(UYFL) UNIV FLORIDA.
XX	
PI	Moyer RW, Li Y, Bawden AL;
DR	WPI; 2002-227161/28.
DR	N-PSDB; ABL56216.
XX	
PT	Novel recombinant entomopox virus vector useful for delivering polynucleotide encoding protein to vertebrate cell, comprises polynucleotide encoding protein operably linked with heterologous promoter sequence. -
XX	
PS	Claim 75; Page 263-266; 326pp; English.
XX	
CC	The invention relates to a recombinant entomopox virus (EPV) vector, comprising a polynucleotide encoding a protein operably linked with a heterologous promoter sequence. The invention also concerns methods for providing gene therapy for genetic deficiency disorders. Vectors of the invention are useful for delivering a polynucleotide encoding a protein to a vertebrate cell preferably a mammalian cell, such as a human cell. The vector is introduced into the vertebrate cell by infection in a viral particle, or by transfection, transduction, or injection either in vitro or in vivo. The vector is useful for the delivery and expression of biologically useful proteins in gene therapy protocols, and for delivering large DNA segments for engineering of vertebrate cells. Polynucleotides of the invention have applications in techniques such as their use as insertion sites for foreign genes of interest, hybridisation probes, for chromosome and gene mapping, in PCR technologies, and in the production of sense or antisense nucleic acids. Vectors of the invention provide for stable integration and expression of heterologous DNA in host cells, and are adapted for accepting large heterologous polynucleotide inserts which can be delivered in an infected or transformed cell and expressed in a stable fraction. The current sequence represents an

XX SQ Sequence 78 AA; 35.6%; Score 42; DB 21; Length 78;
 Query Match 58.3%; Pred. NO. 45;
 Best Local Similarity 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LNMHTYFPNKGCT 13
 |||||
 Db 45 LHPHTYPSKGS 56

RESULT 40
 ABP32719
 ID ABP32719 standard; Protein, 115 AA.
 XX AC ABP32719;
 XX DT 08-JUL-2002 (first entry)
 DT Human ORF1692 protein, SEQ ID NO:3384.

Human, ORF; open reading frame; ORFX; drug screening; diagnosis;
 disease monitoring; cytokine; cell proliferation; cell differentiation;
 immune modulation; haematopoiesis regulation; tissue growth;
 angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 thrombolytic; tumour inhibition; bodily characteristics; fertility;
 behaviour; cancer; proliferative disorder; neurological disorder;
 cardiovascular disease; immune system disorder; organ transplantation;
 tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 vasotrophic; antipsoriatic; antidiabetic; cytosaric; nocrotic;
 neuroprotective; antihypertensive; antidiabetic; thrombolytic;
 cardiant; hypotensive; antitubercular; antiinflammatory; immunomodulator;
 dermatological; analgesic; virucide; antibacterial; fungicide.

XX OS Homo sapiens.
 XX PN WO200190366-A2.
 XX PD 29-NOV-2001.
 XX PF 24-MAY-2001; 2001WO-US17076.
 XX PR 24-MAY-2000; 2000US-206690P.
 XX PA (CURA-) CURAGEN CORP.

Leach MD, Shinkets RA;
 WPI: 2002-106200/14.
 DR N-PSDB; ABN76745.
 DR N-PSDB; ABN76745.
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation
 PT
 XX Claim 10; Page 1105; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,

CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX SQ Sequence 115 AA; 35.6%; Score 42; DB 23; Length 115;
 Query Match 38.1%; Pred. NO. 71;
 Best Local Similarity 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGCTQYTDQIE 21
 |||||
 Db 92 GLNMEGLYRRSGSQTLVEIE 112

Search completed: August 20, 2003, 09:29:33
 Job time : 72.4433 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 23.1649 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-2

Sequence: 1 GLNMHTYFPKGTQYTDQIE 21

Scoring table: BLASTSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	395	3	US-08-856-841-13
2	118	100.0	398	3	US-08-856-841-21
3	118	100.0	415	3	US-08-856-841-20
4	118	100.0	543	3	US-08-856-841-22
5	118	100.0	756	4	US-09-438-268-4
6	44	37.3	119	4	US-09-345-236B-135
7	44	37.3	235	4	US-09-345-236B-108
8	43	36.4	180	3	US-09-153-586-24
9	43	36.4	185	3	US-09-153-586-2
10	43	36.4	378	2	US-08-846-762-87
11	42.5	36.0	151	4	US-09-732-210-1674
12	42.5	36.0	969	4	US-09-252-991A-23580
13	42	35.6	153	4	US-09-732-210-1672
14	42	35.6	157	4	US-09-149-476-350
15	42	35.6	170	4	US-09-732-210-1676
16	42	35.6	584	4	US-09-198-452A-349
17	42	35.6	1501	2	US-08-447-464-3
18	42	35.6	1501	2	US-08-716-679-3
19	42	35.6	1911	1	US-08-348-006B-5
20	42	35.6	1911	1	US-08-800-825A-5
21	42	35.6	1911	5	US-09-158-657-5
22	42	35.6	1911	5	PCT-US94-1016-5
23	41.5	35.2	398	5	US-08-931-220-5
24	41.5	35.2	398	5	PCT-US96-11723-5
25	41.5	35.2	398	5	PCT-US96-0597-1
26	41	34.7	269	3	US-09-252-991A-21483
27	40.5	34.3	16	3	US-08-931-220-2

28	40.5	34.3	16	5	PCT-US95-11723-2	Sequence 2, Appl
29	40.5	34.3	16	5	PCT-US96-0597-15	Sequence 15, Appl
30	40	33.9	187	4	US-09-095-855-203	Sequence 203, Appl
31	40	33.9	187	4	US-09-505-426-203	Sequence 203, Appl
32	40	33.9	465	3	US-09-537-357-52	Sequence 52, Appl
33	40	33.9	1097	4	US-09-252-991A-22579	Sequence 22579, A
34	39	33.1	124	4	US-09-107-532A-4773	Sequence 4773, Ap
35	39	33.1	133	4	US-09-134-001C-4122	Sequence 4122, Ap
36	39	33.1	241	4	US-09-134-001C-5493	Sequence 5493, Ap
37	39	33.1	552	4	US-09-328-352-7229	Sequence 7229, Ap
38	39	33.1	925	1	US-08-252-950D-4	Sequence 4, Appl
39	39	33.1	925	2	US-08-834-108-4	Sequence 4, Appl
40	38	32.2	29	2	US-08-867-087B-52	Sequence 52, Appl
41	38	32.2	137	3	US-09-456-830-7	Sequence 7, Appl
42	38	32.2	137	3	US-09-456-830-15	Sequence 15, Appl
43	38	32.2	137	3	US-09-456-830-19	Sequence 19, Appl
44	38	32.2	137	3	US-09-456-830-37	Sequence 37, Appl
45	38	32.2	137	4	US-09-002-285-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
CLASSIFICATION:
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/214, 658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
JOURNAL: GLUTATHIONE S. TRANSFERASE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
US-08-856-841-13

Query Match 100.0%; Score 118; DB 3; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNNHTYPPNKGTOOYTQIE 21
Db 247 GNNHTYPPNKGTOOYTQIE 267

RESULT 2
US-08-856-841-21
Sequence 21, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214, 658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LTR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 398
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 21:
US-08-856-841-21

Query Match 100.0%; Score 118; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 2,9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
Db 236 GLNMHTYFPNKGTOOYTDQIE 256

RESULT 3

US-08-856-841-20
; Sequence 20, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: N/A
; HYPOTHEICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:

PUBLICATION INFORMATION:

; AUTHORS: COSSART, Y.B.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; AUTHORS: WIDOWS, D.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
; JOURNAL: LANCET
; VOLUME: 1
; ISSUE:
; PAGES: 72 - 73
; DATE: 1975
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 118; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
Db 247 GLNMHTYFPNKGTOOYTDQIE 267

RESULT 4

US-08-856-841-22
; Sequence 22, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543
; TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURES:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOES, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 118; DB 3; Length 543;
Best Local Similarity 100.0%; Pred. No. 4,1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLNMHTYFPNKGTOOYTDOIE 21
DB 395 GLNMHTYFPNKGTOOYTDOIE 415

RESULT 5
US-09-438-268-4
Sequence 4, Application US/09438268
Patent No. 6491907
GENERAL INFORMATION:
APPLICANT: Rabinowitz, Joseph E.
APPLICANT: Samulski, Richard J.
APPLICANT: Xiao, Weidong
TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 5470-186
CURRENT APPLICATION NUMBER: US/09/438,268
CURRENT FILING DATE: 1999-11-10
EARLIER APPLICATION NUMBER: 60/107,840
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/123,651
EARLIER FILING DATE: 1999-03-10
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 756
TYPE: PRT
ORGANISM: Virus
US-09-438-268-4

Query Match 100.0%; Score 118; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 6e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GLNMHTYFPNKGTOOYTDOIE 21

DB 588 GLNMHTYFPNKGTOOYTDOIE 608
RESULT 6
US-09-345-236B-135
Sequence 135, Application US/09345236B
Patent No. 6521454
GENERAL INFORMATION:
APPLICANT: Becnel, James J.
APPLICANT: Tokuo, Fukuda
APPLICANT: Moser, Bettina
APPLICANT: Cockburn, Andrew
APPLICANT: White, Susan E.
APPLICANT: Undeen, Albert H.
TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates
FILE REFERENCE: 21042.0004
CURRENT APPLICATION NUMBER: US/09/345,236B
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 148
SOFTWARE: PatSeq for Windows Version 3.0
SEQ ID NO 135
LENGTH: 119
TYPE: PRT
ORGANISM: mosquitto baculovirus
US-09-345-236B-135

Query Match 37.3%; Score 44; DB 4; Length 119;
Best Local Similarity 50.0%; Pred. No. 7.2;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 6 TYPFNKGTOOYTDOIE 21
DB 88 TCFPSTSTSEYTSWIK 103

RESULT 7
US-09-345-236B-108
Sequence 108, Application US/09345236B
Patent No. 6521454
GENERAL INFORMATION:
APPLICANT: Becnel, James J.
APPLICANT: Tokuo, Fukuda
APPLICANT: Moser, Bettina
APPLICANT: Cockburn, Andrew
APPLICANT: White, Susan E.
APPLICANT: Undeen, Albert H.
TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates
FILE REFERENCE: 21042.0004
CURRENT APPLICATION NUMBER: US/09/345,236B
CURRENT FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 148
SOFTWARE: PatSeq for Windows Version 3.0
SEQ ID NO 108
LENGTH: 235
TYPE: PRT
ORGANISM: mosquitto baculovirus
US-09-345-236B-108

Query Match 37.3%; Score 44; DB 4; Length 235;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 6 TYPFNKGTOOYTDOIE 21
DB 204 TCFPSTSTSEYTSWIK 219

RESULT 8
US-09-153-586-24
Sequence 24, Application US/09153586A
Patent No. 6270772

```

: GENERAL INFORMATION:
: APPLICANT: Burrows et al.
: TITLE OF INVENTION: Recombinant MHC molecules useful for manipulation of
: TITLE OF INVENTION: antigen-specific T-cells
: FILE REFERENCE: 48823
: CURRENT APPLICATION NUMBER: US/09/153,586A
: CURRENT FILING DATE: 1998-09-15
: EARLIER APPLICATION NUMBER: 60/064,552
: EARLIER FILING DATE: 1997-09-16
: EARLIER APPLICATION NUMBER: 60/064,555
: NUMBER OF SEQ. ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ. ID NO 24
: LENGTH: 180
: TYPE: PRT
: ORGANISM: Rattus sp.
: US-09-153-586-24

```

6. Every March
 at Local Similarity
 Matches 7; Conservative 36.4%; Score 43; DB 3; Length 180;
 Conservative 33.3%; Pred. No. 17; Mismatches 4; Indels 2; Gaps 1.

RESULT 9
 US-09-153-586-2
 : Sequence 2, Application US/09153586A
 : Patent NO. 6270772
 : GENERAL INFORMATION:
 : APPLICANT: Burrows et al.
 : TITLE OF INVENTION: Recombinant MHC molecules useful for manipulation of
 : TITLE OF INVENTION: antigen-specific T-cells
 : FILE REFERENCE: 48823
 : CURRENT APPLICATION NUMBER: US/09/153,586A
 : CURRENT FILING DATE: 1998-09-15
 : EARLIER APPLICATION NUMBER: 60/064,552
 : EARLIER FILING DATE: 1997-09-16
 : EARLIER APPLICATION NUMBER: 60/064,555
 : EARLIER FILING DATE: 1997-10-10
 : NUMBER OF SEQ ID NOS: 30
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 185
 : TYPE: PRT
 : ORGANISM: Rattus sp.
 : 09-153-586-2

Query Match	36.4%	Score 43	DB 3	Length 185
Best Local Similarity	33.3%	Pred. No. 18		
Matches	7	Conservative	8	Mismatches 4; Indels 2; Gaps 1
Qy	1	GLANHTTTPENKGTQOYTDQIS	21	
	:	:	: :	
Db	110	GNNYQYVESRG-QETHEFD	128	

```

RESULT 10
US-08-846-762-87
; Sequence 87, Application US/08846762A
; Patent No. 5994072
; GENERAL INFORMATION:
; APPLICANT: Lam, Joseph S.
; APPLICANT: Barrows, Lori
; APPLICANT: Charter, Deborah
; APPLICANT: de Kievit, Teresa
; TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly
; TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa
; FILE REFERENCE: 6580-089
; CURRENT APPLICATION NUMBER: US/08/846,762A

```

```

? CURRENT FILING DATE: 1997-04-30
? NUMBER OF SEQ ID NOS: 100
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 87
? LENGTH: 378
? TYPE: prt
? ORGANISM: Pseudomonas aeruginosa
US-08-646-762-87
```

[illegible]

```

RESULT 11
US-09-732-210-1674
: Sequence 1674, Application US/09732210
: Patent No. 6573361
: GENERAL INFORMATION:
: APPLICANT: Bunkers, Greg J.
: APPLICANT: Liang, Jihong
: APPLICANT: Mitanck, Cindy A.
: APPLICANT: Seale, Jeffrey W.
: APPLICANT: Wu, Yonnie S.
: TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
: FILE REFERENCE: 38-21(1503)B
: CURRENT APPLICATION NUMBER: US/09/732,210
: CURRENT FILING DATE: 2000-12-07
: PRIOR APPLICATION NUMBER: US 60/169,513
: PRIOR FILING DATE: 1999-12-07
: PRIOR APPLICATION NUMBER: US 60/169,340
: PRIOR FILING DATE: 1999-12-07
: NUMBER OF SEQ ID NOS: 1753
: SEQ ID NO 1674
: LENGTH: 151
: TYPE: PR1
: ORGANISM: Mycobacterium tuberculosis
US-09-732-210-1674

```

```

Query Match      36.0%   Score 42.5; DB 4, Length 151;
Best Local Similarity 35.7%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 7; Indels 7; Gaps 1;

QY      1 GLNNHTTFENKGTQOY-----TDQIE 21
          ::|||::|||::|||
Db       53 GRSLDEYFENKRVHQQLIKAPLVTVDRVE 80

```

```

RESULT 12
US-09-252-991A-23580
; Sequence 23580, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23580
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23580

```

Query Match 36.0%; Score 42.5; DB 4; Length 969;
Best Local Similarity 36.4%; Pred. No. 1.5e+02;
Matches 8; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

QY 3 NMHTYFP---NKGTQYTDQIR 21

Db 545 NLHPAPABQSEGYRLTDELE 566

RESULT 13

US-09-732-210-1672

; Sequence 1672; Application US/097322210

; Patent No. 6573361

; GENERAL INFORMATION:

; APPLICANT: Bunkers, Greg J.

; APPLICANT: Liang, Jihong

; APPLICANT: Mittenack, Cindy A.

; APPLICANT: Seale, Jeffrey W.

; APPLICANT: Wu, Yonnie S.

; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use

; FILE REFERENCE: 38-21(15036)B

; CURRENT APPLICATION NUMBER: US/09/732,210

; PRIOR APPLICATION NUMBER: 2000-12-07

; PRIOR FILING DATE: 1999-12-07

; PRIOR APPLICATION NUMBER: US 60/169,513

; PRIOR FILING DATE: 1999-12-07

; PRIOR APPLICATION NUMBER: US 60/169,340

; PRIOR FILING DATE: 1999-12-07

; NUMBER OF SEQ ID NOS: 1753

; SEQ ID NO 1672

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Mycobacterium leprae

US-09-732-210-1672

Query Match 35.6%; Score 42; DB 4; Length 153;

Best Local Similarity 53.3%; Pred. No. 21;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGSTQ 15

Db 55 GRSLEDYFPNKGSTQ 69

RESULT 14

US-09-149-476-350

; Sequence 350; Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: P2002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; PRIOR APPLICATION NUMBER: PCT/US98/04493

; PRIOR FILING DATE: 1998-09-08

; PRIOR APPLICATION NUMBER: 1998-03-06

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,162

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,333

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/038,621

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,626

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,334

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,336

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,163

; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/047,600

; PRIOR FILING DATE: 1997-05-23

; PRIOR APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 35.6%; Score 42; DB 4; Length 157;
Best Local Similarity 36.8%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Oy 2 LNMHTYFPNKGTOGYTDQI 20
Db 62 MNRHEYLKPKMTQSEVDNV 80

RESULT 15
US-09-732-210-1676
Sequence 1676, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mitanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1676
LENGTH: 170
TYPE: PRT
ORGANISM: Streptomyces coelicolor
US-09-732-210-1676

Query Match 35.6%; Score 42; DB 4; Length 170;
Best Local Similarity 42.1%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Oy 1 GLNMHTYFPNKGTOGYTDQ 19
Db 71 GRTLEDFPNKVKVQOEYNE 89

RESULT 16
US-09-198-452A-349

Sequence 349, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 349
LENGTH: 584
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-349

Query Match 35.6%; Score 42; DB 4; Length 584;
Best Local Similarity 77.8%; Pred. No. 98;
tches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

5 HTFPMKGT 13
|||||
Db 573 HAYFPNKET 581

RESULT 17
US-08-447-464-3
Sequence 3, Application US/08447464
Patent No. 5840842
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Yan, Hai
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,464
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/130,570
FILING DATE: 01-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-447-464-3

Query Match 35.6%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 3e+02;
tches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 YFPKGTQY 16
|:|:|:|:|
Db 1057 YWPNRGTEY 1066

RESULT 18
US-08-716-679-3
Sequence 3, Application US/08716679
Patent No. 5846800
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Yan, Hai
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,679
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/130,570
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-716-679-3

Query Match 35.6%; Score 42; DB 2; Length 1501;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 YFPKGTQY 16
|:|:|:|:|
Db 1057 YWPNRGTEY 1066

RESULT 19
US-08-348-006B-5
Sequence 5, Application US/08348006B
Patent No. 5658756
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:

ADDRESSEE: J. MARK HAND
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,006B
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J., MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 189921A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-348-006B-5

Query Match 35.6%; Score 42; DB 1; Length 1911;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 7 YFPNKGTOQY 16
Db 1467 YWPNRGTEY 1476

RESULT 20
US-08-800-825A-5
Sequence 5, Application US/08800825A
Patent No. 5866397
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,825A
FILING DATE: 14-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J. MARK
REGISTRATION NUMBER: 36,545

REFERENCE/DOCKET NUMBER: 189922DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3905
TELEFAX: 732-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-825A-5

Query Match 35.6%; Score 42; DB 2; Length 1911;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 7 YFPNKGTOQY 16
Db 1467 YWPNRGTEY 1476

RESULT 21
US-09-158-657-5
Sequence 5, Application US/09158657
Patent No. 6214564
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A.
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/158,657
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/800,825
FILING DATE: 14-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: HAND, J. MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: 189922DA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3905
TELEFAX: 732-594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-158-657-5

Query Match 35.6%; Score 42; DB 3; Length 1911;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 7 YFPNKGTOQY 16

Db 1467 YWPNRGTEY 1476

RESULT 22

PCT-US94-10166-5
Sequence 5, Application PC/TUS9410166
GENERAL INFORMATION:
APPLICANT: RODAN, GIDEON A
APPLICANT: SCHMIDT, AZRIEL
APPLICANT: RUTLEDGE, SU JANE
TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOHN W. WALLEN III
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10166
FILING DATE: 09-SEPT-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WALLEN, JOHN W III
REGISTRATION NUMBER: 35403
REFERENCE/DOCKET NUMBER: 18992
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1911 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
OS94-10166-5

Query Match 35.6%; Score 42; DB 5; Length 1911;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 7 YFPNKGTOY 16
Db 1467 YWPNRGTEY 1476

RESULT 23
US-08-931-220-5
Sequence 5, Application US/08931220
Patent No. 6030835
GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.
APPLICANT: Kapur M.D., Vivek
TITLE OF INVENTION: Methods and Compositions for Identifying
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEIL, GOSHAL & MANGES
STREET: 2882 Sand Hill Road, Suite 280
CITY: Menlo Park

STATE: CA
COUNTRY: USA
ZIP: 94025-7022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,220
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
APPLICATION NUMBER: US 08/160,965
FILING DATE: 02-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: MGAS 1719
IMMEDIATE SOURCE:
CLONE: SPEB7 (cysteine protease)
PUBLICATION INFORMATION:
AUTHORS: Kapur, V.
AUTHORS: Topouzis, S.
AUTHORS: Majesky, M. W.
AUTHORS: Li, L.-L.
AUTHORS: Hamrick, M. R.
AUTHORS: Hamill, R. J.
AUTHORS: Patil, J. M.
AUTHORS: Musser, J. M.
TITLE: A conserved Streptococcus pyogenes
TITLE: extracellular cysteine protease cleaves human
TITLE: fibronectin and degrades vitronectin
JOURNAL: Microb. Pathog.
VOLUME: 15
PAGES: 327-346
DATE: 1993

US-08-931-220-5

Query Match 35.2%; Score 41.5; DB 3; Length 398;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

OY 2 LNMHTYFPNKGTOY 17
Db 201 MKHTNY-PNKGKDYT 215

RESULT 24
PCT-US95-11723-5
Sequence 5, Application PC/TUS9511723
GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.
APPLICANT: Kapur M.D., Vivek
TITLE OF INVENTION: Methods and Compositions for

TITLE OF INVENTION: Identifying Streptococcus
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAW OFFICES OF BARBARA RAE-VENTER
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11723
FILING DATE: 14-SEP-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/03US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6205
TELEFAX: (415) 424-8760
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: MGAS 1719
IMMEDIATE SOURCE:
CLONE: SPEB7 (cysteine protease)
PUBLICATION INFORMATION:
AUTHORS: Kapur, V.
AUTHORS: Topouzis, S.
AUTHORS: Majesky, M. W.
AUTHORS: Li, L., L.
AUTHORS: Hamrick, M. R.
AUTHORS: Hamill, R. J.
AUTHORS: Patel, J. M.
AUTHORS: Musser, J. M.
TITLE: A conserved Streptococcus pyogenes
TITLE: extracellular cysteine protease cleaves human
TITLE: fibronectin and degrades vitronectin
JOURNAL: Microb. Pathog.
VOLUME: 15
PAGES: 327-346
DATE: 1993
PCT-US95-11723-5

Query March 35.2%, Score 41.5, DB 5, Length 398;
Best Local Similarity 50.0%, Pred. No. 76;
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

CY 2 LNMHTYFPNKGTOOYT 17
Db 201 MKYHNY-PNKGKLDY 215

RESULT 25
PCT-US96-05997-1
Sequence 1, Application PC/TUS9605997

GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.
APPLICANT: Kapur M.D., Vivek
APPLICANT: Ananthaswamy, H. N.
APPLICANT: Fernandez, A.
TITLE OF INVENTION: Use of extracellular cysteine protease
TITLE OF INVENTION: to inhibit cell proliferation
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEIL, GOTSCHAL & MANGES
STREET: 2882 Sand Hill Road, Suite 280
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-7022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05997
FILING DATE: 01-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,542
FILING DATE: 14-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/160,965
FILING DATE: 02-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter Ph.D., Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: BAYL-004/02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: MGAS 1719
IMMEDIATE SOURCE:
CLONE: SPEB7
PCT-US96-05997-1

Query March 35.2%, Score 41.5, DB 5, Length 398;
Best Local Similarity 50.0%, Pred. No. 76;
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

CY 2 LNMHTYFPNKGTOOYT 17
Db 201 MKYHNY-PNKGKLDY 215

RESULT 26
US-09-252-991A-21483
Sequence 21483, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21483
LENGTH: 269
TYPE: prt
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21483

Query Match 34.7%; Score 41; DB 4; Length 269;
Best Local Similarity 43.8%; Pred. No. 58;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HTYFPNKGTOOYTQ 19
Db 100 VTYIDASGNRYVTDQ 115

RESULT 27

US-08-931-220-2
Sequence 2, Application US/08931220
Patent No. 6030835

GENERAL INFORMATION:
APPLICANT: Musser M.D., James M.

APPLICANT: Kapur M.D., Vivek

TITLE OF INVENTION: Methods and Compositions for Identifying

TITLE OF INVENTION: Group A Streptococcus

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: WEIL, GOTSHAL & MANGES

STREET: 2882 Sand Hill Road, Suite 280

CITY: Menlo Park

STATE: CA

COUNTRY: USA

ZIP: 94025-7022

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/931,220

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/306,542

FILING DATE: 14-SEP-1994

APPLICATION NUMBER: US 08/160,965

FILING DATE: 02-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter Ph.D., Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: BAYL-004/01US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 926-6200

TELEFAX: (415) 954-3713

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

ORGANISM: Streptococcus pyogenes

POSITION IN GENOME:

MAP POSITION: 203

US-08-931-220-2

Query Match 34.3%; Score 40.5; DB 3; Length 16;
Best Local Similarity 61.5%; Pred. No. 2.6;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 5 HTYFPNKGTOOYT 17
Db 2 HNY-PNKGKLDY 13

RESULT 28

PCT-US95-11723-2
Sequence 2, Application PC/TUS9511723

GENERAL INFORMATION:

APPLICANT: Musser M.D., James M.

APPLICANT: Kapur M.D., Vivek

TITLE OF INVENTION: Methods and Compositions for

TITLE OF INVENTION: Identifying Streptococcus

NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAW OFFICES OF BARBARA RAE-VENTER

STREET: P.O. Box 60039

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/11723

FILING DATE: 14-SEP-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/306,542

FILING DATE: 14-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter Ph.D., Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: BAYL-004/03US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 926-6205

TELEFAX: (415) 424-8760

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

ORGANISM: Streptococcus pyogenes

POSITION IN GENOME:

MAP POSITION: 203

PCT-US95-11723-2

Query Match 34.3%; Score 40.5; DB 5; Length 16;
Best Local Similarity 61.5%; Pred. No. 2.6;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 5 HTYFPNKGTOOYT 17
Db 2 HNY-PNKGKLDY 13

RESULT 29

PCT-US96-05997-15

```

; Sequence 15 Application PC/TUS9605997
; GENERAL INFORMATION:
; APPLICANT: Musser M.D., James M.
; APPLICANT: Kapur M.D., Vivek
; APPLICANT: Ananthaswamy, H. N.
; APPLICANT: Fernandez, A.
; TITLE OF INVENTION: Use of extracellular cysteine protease
; TITLE OF INVENTION: to inhibit cell proliferation
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WEIL, GOTSHAL & MANGES
; STREET: 2882 Sand Hill Road, Suite 280
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025-7022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05997
; FILING DATE: 01-MAY-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,542
; FILING DATE: 14-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/160,965
; FILING DATE: 02-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Ventler Ph.D., Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: BAYL-004/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 926-6200
; TELEFAX: (415) 854-3713
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; POSITION IN GENOME:
; MAP POSITION: 203
PCT-US96-05997-15
Query Match 34.3%; Score 40.5; DB 5; Length 16;
Best Local Similarity 61.5%; Pred. No. 2.6;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 5 HTYPEPNKGTQOYT 17
Db 2 HNY-PNKGDKOYT 13

RESULT 30
US-09-095-955-203
; Sequence 203; Application US/0905855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for

```

```

1 TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
2
3 NUMBER OF SEQUENCES: 208
4
5 CORRESPONDENCE ADDRESS:
6
7 ADDRESSEE: Law Offices of Ann W. Speckman
8 STREET: 2601 Elliott Avenue, Suite 4185
9 CITY: Seattle
10 STATE: WA
11
12 COUNTRY: USA
13
14 ZIP: 98121
15
16 COMPUTER READABLE FORM:
17
18 MEDIUM TYPE: Diskette
19
20 COMPUTER: IBM Compatible
21
22 OPERATING SYSTEM: DOS
23
24 SOFTWARE: PaeCSQ for Windows Version 2.0
25
26 CURRENT APPLICATION DATA:
27
28 APPLICATION NUMBER: US/09/095,855
29
30 FILING DATE:
31
32 CLASSIFICATION:
33
34 PRIOR APPLICATION DATA:
35
36 APPLICATION NUMBER: 08/705,347
37
38 FILING DATE: 29-AUG-1996
39
40 APPLICATION NUMBER: 08/873,970
41
42 FILING DATE: 12-JUN-1997
43
44 APPLICATION NUMBER: 08/997,362
45
46 FILING DATE: 23-DEC-1997
47
48 ATTORNEY/AGENT INFORMATION:
49
50 NAME: Sleath, Janet
51
52 REGISTRATION NUMBER: 37,007
53
54 REFERENCE/DOCKET NUMBER: 11000.1002c3
55
56 TELECOMMUNICATION INFORMATION:
57
58 TELEPHONE: 206-269-0565
59
60 TELEFAX: 206-269-0563
61
62 TELEX:
63
64 INFORMATION FOR SEQ ID NO: 203:
65
66 SEQUENCE CHARACTERISTICS:
67
68 LENGTH: 187 amino acids
69
70 TYPE: amino acid
71
72 STRANDEDNESS: single
73
74 TOPOLOGY: linear
75
76 MOLECULE TYPE: protein
77
78 US-09-095-855-203
79
80
81 Query Match 33.9%; Score 40; DB 3; Length 187;
82 Best Local Similarity 58.3%; Pred. No. 55;
83 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
84
85 QY 10 NKGTQQTDDTE 21
86
87 DB 161 DKSTHYTNQID 172
88
89
90 RESULT 31
91
92 US-09-205-426-203
93
94 Sequence 203, Application US/09205426
95
96 Patent No. 6406704
97
98 GENERAL INFORMATION:
99
100 APPLICANT: Watson, James D.
101
102 APPLICANT: Tan, Paul L. J.
103
104 TITLE OF INVENTION: Compounds and Methods for Treatment and
105
106 TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
107
108 FILE REFERENCE: 11000.1002c4
109
110 CURRENT APPLICATION NUMBER: US/09/205,426
111
112 CURRENT FILING DATE: 1998-12-04
113
114 EARLIER APPLICATION NUMBER: 09/095,855
115
116 EARLIER FILING DATE: 1998-06-11
117
118 EARLIER APPLICATION NUMBER: 08/997,362
119
120 EARLIER FILING DATE: 1997-12-23
121
122 EARLIER APPLICATION NUMBER: 08/873,970
123
124 EARLIER FILING DATE: 1997-06-12
125
126 EARLIER APPLICATION NUMBER: 08/705,347
127
128 EARLIER FILING DATE: 1996-08-29
129
130 NUMBER OF SEQ ID NOS: 208
131
132 SOFTWARE: FastSeq for Windows Version 3.0

```

SEQ ID NO 203
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186) ... (186)
US-09-205-426-203

Query Match 33.9%; Score 40; DB 4; Length 187;
Best Local Similarity 58.3%; Pred. No. 55;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 NKGTQOYTDQ 21
DB 161 DKSTHQYTNQID 172

RESULT 32
US-537-357-52
Sequence 52, Application US/09537357
Patent No. 6271018
GENERAL INFORMATION:
APPLICANT: Alan Brash
APPLICANT: Nathalie Tjjet
TITLE OF INVENTION: MUSKELON (CUCUMIS MELO) HYDROPEROXIDE
FILE REFERENCE: 06027.0002
CURRENT APPLICATION NUMBER: US/09/537,357
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 52
LENGTH: 465
TYPE: PRT
ORGANISM: Cucumis melo
US-09-537-357-52

Query Match 33.9%; Score 40; DB 3; Length 465;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 HTPPNKGTQOY 16
DB 89 HDYFYNQGREY 100

RESULT 33
US-252-991A-22579
Sequence 22579, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,768
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22579
LENGTH: 1097
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22579

Query Match 33.9%; Score 40; DB 4; Length 1097;
Best Local Similarity 46.7%; Pred. No. 4.3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 6 TYPFNKGTQOYTDQ 20
DB 769 TEPFRRGQOQGGYQ 783

RESULT 34
US-09-107-532A-4773
Sequence 4773, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4773:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...124
SEQUENCE DESCRIPTION: SEQ ID NO: 4773:
US-09-107-532A-4773

Query Match 33.1%; Score 39; DB 4; Length 124;
Best Local Similarity 58.3%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 FPNNKGTQOYTDQ 19
DB 64 FPNNGVQGYYSQ 75

RESULT 35
US-09-134-001C-4122
Sequence 4122, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4122
LENGTH: 133
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4122

Query Match 33.1%; Score 39; DB 4; Length 133;
Best Local Similarity 38.9%; Pred. No. 53;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 MHTYFPNKGTOOYTDQIE 21
112 LSTYVVKKGKDYDTSKIK 129

RESULT 36
US-09-134-001C-5493
Sequence 5493, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5493
LENGTH: 241
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5493

Query Match 33.1%; Score 39; DB 4; Length 241;
Best Local Similarity 63.6%; Pred. No. 1,1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 YFPNKGTOOYT 17
223 YFNGKQKQRTT 233

RESULT 37
US-09-328-352-7229
Sequence 7229, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7229
LENGTH: 532
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7229

Query Match 33.1%; Score 39; DB 4; Length 532;

Best Local Similarity 53.3%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 2 LNMHTYFPNKGTOOY 16
Db 480 LNSSEPNNTOOY 494

RESULT 38
US-08-252-995D-4
Sequence 4, Application US/08252995D
Patent No. 5650501
GENERAL INFORMATION:
APPLICANT: Dennis, James W
APPLICANT: Heffernan, Mike
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERSKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,995D
FILING DATE: 02-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurdzyk, Linda M
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 925 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-252-995D-4

Query Match 33.1%; Score 39; DB 1; Length 925;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 HTYFPNKGTOOYTD 18
Db 460 HTVSPNRDPQDYD 473

RESULT 39
US-08-834-108-4
Sequence 4, Application US/08834108
Patent No. 5976893
GENERAL INFORMATION:
APPLICANT: Dennis, James W
APPLICANT: Heffernan, Mike
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERSKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario

COUNTRY: Canada
ZIP: MSH 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,108
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurd/dyk, Linda M
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 925 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-834-108-4

Query Match 33.1% Score 39; DB 2; Length 925;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
Qy 5 HTYPNKGTOOYTD 18
DB 460 HTSPNRFQOYDPD 473

RESULT 40
US-08-867-087B-52
Sequence 52, Application US/08867087B
Patent No. 5990386
GENERAL INFORMATION:
APPLICANT: An, Gynheung
TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
TITLE OF SEQUENCES: AND APICAL DOMINANCE IN PLANTS
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh &
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
CITY: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,087B
FILING DATE: June 2, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/323,449
FILING DATE: October 14, 1994
APPLICATION NUMBER: U.S. 08/485,981
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan, E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 4630-47071
TELECOMMUNICATION INFORMATION:

TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acid residues
TYPE: amino acid
TOPOLOGY: linear
US-08-867-087B-52

Query Match 32.2% Score 38; DB 2; Length 29;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 3 NMHTYFPNKG 12
DB 4 NMHTYFPNKG 13

Search completed: August 20, 2003, 09:34:26
Job time : 24.1649 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comogen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:30:58 ; Search time 85.299 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-2

Perfect score: 118
Sequence: 1 GINMHTYPPNKGTOOYTDOIE 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 497079 seqs, 131961718 residues

1 number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	100.0	756	US-10-205-942-4	Sequence 4, Appli
2	45	38.1	225	US-09-895-913A-86	Sequence 86, Appli
3	43.5	36.9	437	US-10-156-761-10200	Sequence 10200, A
4	43	36.4	180	US-09-858-580-24	Sequence 24, Appli
5	43	36.4	180	US-09-847-172-24	Sequence 24, Appli
6	43	36.4	185	US-09-858-580-2	Sequence 2, Appli
7	43	36.4	185	US-09-847-172-2	Sequence 2, Appli
8	43	36.4	297	US-10-205-823-70	Sequence 70, Appli
9	43	36.4	378	US-10-216-209-87	Sequence 87, Appli
10	42	35.6	157	US-09-809-391-350	Sequence 350, App
11	42	35.6	173	US-10-156-761-12492	Sequence 12492, A
12	42	35.6	283	US-09-815-242-13476	Sequence 13476, A
13	42	35.6	436	US-09-815-242-13218	Sequence 13218, A
14	42	35.6	1502	US-09-808-602-54	Sequence 54, Appli
15	42	35.6	1502	US-09-800-198-44	Sequence 44, Appli

16	42	35.6	1948	US-09-808-602-55	Sequence 55, Appli
17	42	35.6	1948	US-09-800-198-45	Sequence 45, Appli
18	41.5	35.2	248	US-10-002-784A-25	Sequence 25, Appli
19	41.5	35.2	398	US-10-002-784A-24	Sequence 24, Appli
20	41.5	35.2	468	US-10-002-784A-27	Sequence 27, Appli
21	40.5	34.3	642	US-10-024-632-6	Sequence 6, Appli
22	40	33.9	182	US-09-738-626-4147	Sequence 4147, Ap
23	40	33.9	187	US-10-205-979-49	Sequence 49, Appli
24	40	33.9	187	US-10-051-643-203	Sequence 203, App
25	40	33.9	254	US-09-815-242-5434	Sequence 5434, Ap
26	40	33.9	254	US-09-815-242-12274	Sequence 12274, A
27	40	33.9	465	US-09-884-260A-52	Sequence 52, Appli
28	40	33.9	565	US-10-128-714-8030	Sequence 8080, Ap
29	40	33.9	654	US-09-815-242-10361	Sequence 10361, A
30	40	33.9	869	US-10-190-435-4	Sequence 4, Appli
31	40	33.9	1244	US-09-815-915-8	Sequence 8, Appli
32	40	33.9	1507	US-09-924-154-17	Sequence 17, Appli
33	39	33.1	107	US-10-156-761-10289	Sequence 10289, A
34	39	33.1	251	US-10-156-761-9227	Sequence 9227, Ap
35	39	33.1	324	US-09-905-290A-4	Sequence 4, Appli
36	39	33.1	380	US-09-738-626-6372	Sequence 6372, Ap
37	39	33.1	424	US-09-738-626-4029	Sequence 4029, Ap
38	39	33.1	605	US-10-238-075-691	Sequence 691, App
39	39	33.1	713	US-10-116-519-8	Sequence 8, Appli
40	39	33.1	1645	US-10-156-761-12609	Sequence 12609, A
41	38.5	32.6	543	US-10-156-761-13485	Sequence 13485, A
42	38.5	32.6	547	US-08-834-705-2	Sequence 2, Appli
43	38.5	32.6	1158	US-09-815-242-4594	Sequence 4594, Ap
44	38.5	32.6	1207	US-09-815-242-10626	Sequence 10626, A
45	38	32.2	137	US-10-099-285-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-205-942-4
; Sequence 4, Application US/10205942
; Publication No. US20030053990A1
GENERAL INFORMATION:
; APPLICANT: University of No. US20030053990A1th Carolina-Chapel Hill
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/10/205,942
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Adeno-associated virus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(2271)
; OTHER INFORMATION: B19/AAV chimeric capsid coding sequence
US-10-205-942-4
Query Match 100.0%; Score 118; DB 15; Length 756;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GINMHTYPPNKGTOOYTDOIE 21
|||||
Db 588 GINMHTYPPNKGTOOYTDOIE 608
|||||
RESULT 2
US-09-895-913A-86
; Sequence 86, Application US/09895913A
; Patent No. US20020160456A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Kleinhous, Harold
/ APPLICANT: Al-Garawi, Amal
/ APPLICANT: Miller, Charles
/ APPLICANT: Tomb, Jean Francois
/ APPLICANT: Oomen, Raymond P.
/ TITLE OF INVENTION: Identification of Polynucleotides
/ TITLE OF INVENTION: Encoding No. US20020160456a1 Helicobacter Polypeptides in the
/ FILE REFERENCE: 06133/043002
/ CURRENT APPLICATION NUMBER: US/09/895,913A
/ CURRENT FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: US 08/881,227
/ PRIOR FILING DATE: 1997-06-24
/ NUMBER OF SEQ ID NOS: 368
/ SOFTWARE: PasteSeq for Windows Version 4.0
/ SEQ ID NO 86
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Helicobacter pylori
/ 3-895-913A-86

Query Match          36.1%; Score 45; DB 10; Length 225;
Best Local Similarity 47.4%; Pred. No. 25;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQ 19
DB 173 GLSLNAVFNVFNQYTDQ 191

RESULT 3
US-10-156-761-10200
/ Sequence 10200, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 10200
/ LENGTH: 437
/ TYPE: PRT
/ ORGANISM: Streptomyces avermitilis
/ US-10-156-761-10200

Query Match          36.9%; Score 43.5; DB 15; Length 437;
Best Local Similarity 48.0%; Pred. No. 89;
Matches 12; Conservative 1; Mismatches 5; Indels 7; Gaps 2;

QY 1 GLNMTYFPNKGTOQYTDQ 19
DB 268 GNM-HTSLPADAAFPNPHAGYTDQ 291

RESULT 4
US-09-858-580-24
/ Sequence 24, Application US/09858580
/ Patent No. US20020176864A1
/ GENERAL INFORMATION:
/ APPLICANT: Burrows et al.
```

```
/ TITLE OF INVENTION: Recombinant MHC molecules useful for manipulation of
/ TITLE OF INVENTION: antigen-specific T-Cells
/ FILE REFERENCE: 48823
/ CURRENT APPLICATION NUMBER: US/09/858,580
/ CURRENT FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: 09/153,586
/ PRIOR FILING DATE: 1998-09-15
/ PRIOR APPLICATION NUMBER: 60/064,552
/ PRIOR FILING DATE: 1997-09-16
/ PRIOR APPLICATION NUMBER: 60/064,555
/ PRIOR FILING DATE: 1997-10-10
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 24
/ LENGTH: 180
/ TYPE: PRT
/ ORGANISM: Rattus sp.
/ US-09-858-580-24

Query Match          36.4%; Score 43; DB 10; Length 180;
Best Local Similarity 33.3%; Pred. No. 40;
Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 1 GLNMTYFPNKGTOQYTDQIE 21
DB 105 GNMVQYVESRG--QTFHEFD 123

RESULT 5
US-09-847-172-24
/ Sequence 24, Application US/09847172
/ Publication No. US20030007978A1
/ GENERAL INFORMATION:
/ APPLICANT: OREGON HEALTH AND SCIENCES UNIVERSITY
/ APPLICANT: BURROWS, GREGORY G.
/ APPLICANT: VANDENBARK, ARTHUR A.
/ TITLE OF INVENTION: RECOMBINANT MHC MOLECULES USEFUL FOR MANIPULATION OF ANTIGEN-SPEC
/ FILE REFERENCE: 899-58137
/ CURRENT APPLICATION NUMBER: US/09/847,172
/ CURRENT FILING DATE: 2001-05-01
/ PRIOR APPLICATION NUMBER: US 60/200,942
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 09/153,586
/ PRIOR FILING DATE: 1998-09-15
/ PRIOR APPLICATION NUMBER: US 60/064,555
/ PRIOR FILING DATE: 1997-10-10
/ PRIOR APPLICATION NUMBER: US 60/064,552
/ PRIOR FILING DATE: 1997-09-16
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 24
/ LENGTH: 180
/ TYPE: PRT
/ ORGANISM: Rattus sp.
/ US-09-847-172-24

Query Match          36.4%; Score 43; DB 11; Length 180;
Best Local Similarity 33.3%; Pred. No. 40;
Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 1 GLNMTYFPNKGTOQYTDQIE 21
DB 105 GNMVQYVESRG--QTFHEFD 123

RESULT 6
US-09-858-580-2
/ Sequence 2, Application US/09858580
/ Patent No. US20020176864A1
/ GENERAL INFORMATION:
/ APPLICANT: Burrows et al.
/ TITLE OF INVENTION: Recombinant MHC molecules useful for manipulation of
```

Query Match	36.4%	Score 43	DB 15	Length 378
Best Local Similarity	38.9%	Pred. No. 91		
Matches	7	Conservative	6	Mismatches 5
				Indels 0
				Gaps 0
OY	3	NMHTYPPNKGTOYTTDOI	20	
		: : :	: :	

Db 125 MNNSPPEEGNRQLTSKL 142

RESULT 10

US-09-809-391-350
; Sequence 350, Application US/09809391
; Publication No. US2003049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 350
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (157)
; OTHER INFORMATION: Xaa equals stop translation
US-09-809-391-350

Query Match

35.6%; Score 42; DB 11; Length 157;
Best Local Similarity 36.8%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 2 LNMTYFPNKGTOQYTDQI 20
DB 62 MNHRYLPKMPDQSEVDNV 80

RESULT 11

US-10-156-761-12492
; Sequence 12492, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIRAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12492
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12492

Query Match

35.6%; Score 42; DB 15; Length 173;
Best Local Similarity 42.1%; Pred. No. 55;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQI 19
DB 74 GRTLEDYFPNKGTVQGEVNE 92

RESULT 12

US-09-815-242-13476

; Sequence 13476, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13476
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13476

QY 4 MHTYFPNKGTOQYTDQI 20
DB 203 LFTSVNKNKSQYTERV 219

Query Match
35.6%; Score 42; DB 9; Length 283;
Best Local Similarity 41.2%; Pred. No. 94;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 MHTYFPNKGTOQYTDQI 20
DB 203 LFTSVNKNKSQYTERV 219

RESULT 13

US-09-815-242-13218
; Sequence 13218, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

```

1 PRIOR APPLICATION NUMBER: 60/253,625
2 PRIOR FILING DATE: 2000-11-27
3 PRIOR APPLICATION NUMBER: 60/257,931
4 PRIOR FILING DATE: 2000-12-22
5 PRIOR APPLICATION NUMBER: 60/269,308
6 PRIOR FILING DATE: 2001-02-16
7 NUMBER OF SEQ ID NOS: 14110
8 SOFTWARE: FastSeq for Windows Version 4.0
9 SEQ ID NO 13218
10 LENGTH: 436
11 TYPE: prt
12 ORGANISM: Streptococcus pneumoniae
13
14 US-09-815-242-13218

```

Query Match	35.6%;	Score 42;	DB 9;	Length 436;
Best Local Similarity	35.0%;	Pred. No. 1.5e+02;		
Matches	7;	Conservative	5;	Mismatches 8;
				Indels 0;
				Gaps 0;

1 GLNMHTYFPNNKGTQQYTDQI 20
347 GAYIHGVRPCKGTEEYMSKL 366

RESULT 14
 US-09-808-602-54
 ; Sequence 54, Application US/09808602
 ; Patent No.: US2002015115A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vernet, Corine A
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Herrman, John L
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: MacDougall, John
 ; TITLE OF INVENTION: No. US2002015115A1 Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 15966-697 CIP
 ; CURRENT APPLICATION NUMBER: US/09/808,602
 ; CURRENT FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 09/800,198
 ; PRIOR FILING DATE: 2001-03-05
 ; PRIOR APPLICATION NUMBER: 60/186,596
 ; PRIOR FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 54
 ; LENGTH: 1502
 ; TYPE: prt
 ORGANISM: Homo sapiens
 US-09-808-602-54

Query Match	35.6%	Score 42;	DB 10;	Length 1502;
Best Local Similarity	60.0%	Pred. No. 6e+02;		
Matches	6;	Conservative	3;	Mismatches 1;
				Indels 0;
				Gaps 0

```
QY      7 YFPNKGTOQY 16
         |:||:|:|:|
Db      1058 YWPNRGTEIY 1067
```

RESULT 15
US-09-800-198-44
; Sequence 44, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Hermann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S

```

APPLICANT: Rastelli, Luca
TITLE OF INVENTION: NOVEL PROTEIN AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-697
CURRENT APPLICATION NUMBER: US/09/800,198
CURRENT FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44
LENGTH: 1502
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-198-44

```

Query Match	35.6%	Score 42;	DB 11;	Length 1502;
Best Local Similarity	60.0%	Pred. No. 6e+02;		
Matches	6;	Conservative	3;	Mismatches 1; Indels 0; Gaps 0;

```
QY      7 YFPNKGTOQY 16
         |::|::|:
Db      1058 YWPNRGTEY 1067
```

```

RESULT 16
US-09-808-602-55
; Sequence 55, Application US/09808602
; Patent No. US2002015115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: NO. US2002015115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808, 602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800, 198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186, 596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 1948
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-55

```

Query Match	35.6%	Score 42;	DB 10;	Length 1948;
Best Local Similarity	60.0%	Pred. No. 8e+02;		
Matches	6;	Conservative	3;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY	7	YFPNKGTOQY	16
		: : : :	
Db	1504	YWPNRGTETY	1513

RESULT 17
US-09-800-198-45
Sequence 45: Application US/09800199
Publication No. US20030087816A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corrie AM
APPLICANT: Hernandez, Elina
APPLICANT: Shimkets, Richard A
APPLICANT: Herxmann, John L
APPLICANT: Majumder, Kunund
APPLICANT: Mishra, Vishna

APPLICANT: Mezes, Peter S
APPLICANT: Rastelli, Luca
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-697
CURRENT APPLICATION NUMBER: US/09/800,198
CURRENT FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 45
LENGTH: 1948
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-198-45

Query Match 35.6%; Score 42; DB 11; Length 1948;
Best Local Similarity 60.0%; Pred. No. 8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

7 YFPNKGTOOY 16
|:|:|:|:|
Db 1504 YFPNKGTOOY 1513

RESULT 18
US-10-002-784A-25
Sequence 25, Application US/10002784A
Publication No. US20030036644A1
GENERAL INFORMATION:
/33
APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/233/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 25
LENGTH: 248
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: mutant streptococcal pyrogenic exotoxin-B
US-10-002-784A-25

Query Match 35.2%; Score 41.5; DB 15; Length 248;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 LNMHTYFPNKGTOOYT 17
:|:|:|:|:|
Db 56 MKYHNY-PNKGKLDY 70

RESULT 19
US-10-002-784A-24
Sequence 24, Application US/10002784A
Publication No. US20030036644A1
GENERAL INFORMATION:
/33
APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/233/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 24

LENGTH: 398
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: mutant streptococcal pyrogenic exotoxin B prosegment
US-10-002-784A-24

Query Match 35.2%; Score 41.5; DB 15; Length 398;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 LNMHTYFPNKGTOOYT 17
:|:|:|:|:|
Db 201 MKYHNY-PNKGKLDY 215

RESULT 20
US-10-002-784A-27
Sequence 27, Application US/10002784A
Publication No. US20030036644A1
GENERAL INFORMATION:
/33
APPLICANT: Ulrich, Robert G.
TITLE OF INVENTION: Bacterial Superantigen Vaccines
FILE REFERENCE: 003/233/SAP
CURRENT APPLICATION NUMBER: US/10/002,784A
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 08/882,431; 09/144,776
PRIOR FILING DATE: 97-06-25; 98-09-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 27
LENGTH: 468
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: mutant SpeA-mutant SpeB fusion
US-10-002-784A-27

Query Match 35.2%; Score 41.5; DB 15; Length 468;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 LNMHTYFPNKGTOOYT 17
:|:|:|:|:|
Db 276 MKYHNY-PNKGKLDY 290

RESULT 21
US-10-024-632-6
Sequence 6, Application US/10024632
Publication No. US20020170093A1
GENERAL INFORMATION:
APPLICANT: Monsanto Technology LLC
APPLICANT: He, Steve S.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION
FILE REFERENCE: 38-21(51837)B
CURRENT APPLICATION NUMBER: US/10/024,632
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,896
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 642
TYPE: PRT
ORGANISM: Oryza sativa
US-10-024-632-6

Query Match 34.3%; Score 40.5; DB 14; Length 642;
Best Local Similarity 42.9%; Pred. No. 4e+02;

Matches 9; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
 Qy 1 GNNMHTYFPNKGTOYTDQIE 21
 ||:|||||:
 Db 349 GUSTHINFP---LENYRDEIE 366

RESULT 22
 US-09-738-626-4147
 ; Sequence 4147, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4147
 ; LENGTH: 182
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-4147

Query Match 33.9%; Score 40; DB 10; Length 182;
 Best Local Similarity 53.3%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 GNNMHTYFPNKGTOQ 15
 ||:|||||:
 Db 84 GRTLEHYFPNKLHQQ 98
 RESULT 23
 US-10-205-979-49
 ; Sequence 49, Application US/10205979
 ; Publication No. US20030147861A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Tan, Paul L. J.
 ; APPLICANT: Abernethy, Nevyn
 ; TITLE OF INVENTION: Compounds and Methods for the Modulation
 ; FILE REFERENCE: 11000.1063U
 ; CURRENT APPLICATION NUMBER: US/10/205,979
 ; CURRENT FILING DATE: 2002-07-25
 ; PRIOR APPLICATION NUMBER: 60/308,446
 ; PRIOR FILING DATE: 2001-07-26
 ; NUMBER OF SEQ ID NOS: 52
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 49
 ; LENGTH: 187
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium vaccae
 ; FEATURE: VARIANT
 ; NAME/KEY: (1)...(187)
 ; LOCATION: (1)...(187)

OTHER INFORMATION: Xaa = Any Amino Acid
 US-10-205-979-49

Query Match 33.9%; Score 40; DB 12; Length 187;
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 10 NKGTQOYTDQIE 21
 :|||||:
 Db 161 DKSTHQYTNQID 172

RESULT 24
 US-10-051-643-203
 ; Sequence 203, Application US/10051643
 ; Publication No. US20020197265A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Tan, Paul L. J.
 ; TITLE OF INVENTION: Methods and Compounds for the Treatment
 ; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
 ; FILE REFERENCE: 11000.1008C2
 ; CURRENT APPLICATION NUMBER: US/10/051,643
 ; CURRENT FILING DATE: 2002-01-18
 ; PRIOR APPLICATION NUMBER: US09/156,181
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: US 08/996,624
 ; PRIOR FILING DATE: 1997-12-23
 ; NUMBER OF SEQ ID NOS: 208
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 203
 ; LENGTH: 187
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium vaccae
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (186)...(186)
 ; US-10-051-643-203

Query Match 33.9%; Score 40; DB 14; Length 187;
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 10 NKGTQOYTDQIE 21
 :|||||:
 Db 161 DKSTHQYTNQID 172

RESULT 25
 US-09-815-242-5434
 ; Sequence 5434, Application US/09815242
 ; Patent No. US2002061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zykend, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5434
LENGTH: 254
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5434

Query Match 33.9%; Score 40; DB 9; Length 254;
Best Local Similarity 37.5%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

1 GUNMHTY----FNNKGTQOYTDQI 20
DB 98 GINDSRYLLEFPSNEVPHYTDQL 121

RESULT 26
US-09-815-242-12274
Sequence 12274, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12274
LENGTH: 254
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12274

Query Match 33.9%; Score 40; DB 9; Length 254;
Best Local Similarity 37.5%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

1 GUNMHTY----FNNKGTQOYTDQI 20
DB 98 GINDSRYLLEFPSNEVPHYTDQL 121

RESULT 27
US-09-884-260A-52
Sequence 52, Application US/09884260A
Patent No. US20020098570A1
GENERAL INFORMATION:
APPLICANT: Alan Braash
APPLICANT: Nathalie Tijer
TITLE OF INVENTION: MUSKMELOX (CUCUMIS MELO) HYDROPEROXIDE
TITLE OF INVENTION: LYSASE AND USES THEREOF
FILE REFERENCE: 06027.000202
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/537,357
PRIOR FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 52
LENGTH: 465
TYPE: PRT
ORGANISM: Cucumis melo
US-09-884-260A-52

Query Match 33.9%; Score 40; DB 9; Length 465;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

5 HTYFPNKG 16
DB 89 HDYFYNQGREY 100

RESULT 28
US-10-128-714-8080
Sequence 8080, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Mengqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshekin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT FILING DATE: 2002-04-23
CURRENT FILING DATE: 2002-04-23
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/316,362
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: Patent version 3.1
SEQ ID NO 8080
LENGTH: 565
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-8080

Query Match 33.9%; Score 40; DB 15; Length 565;
Best Local Similarity 75.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

5 HTYFPNKG 12
DB 21 HTYVPRNG 28


```

RESULT 29
US-09-815-242-10361
; Sequence 10361, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyckind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10361
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10361

Query Match          33.9%; Score 40; DB 9; Length 654;
Blast Local Similarity 45.0%; Pred. No. 4.9e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1

OY      1  GLNHHYTPNK--GTOQYTD 18
      |||: : ||| : |||:
366  GLNIGVQWANKVNGAQDTN 385

RESULT 30
US-10-190-435-4
; Sequence 4, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Esterlitta J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PPI8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Description of Artificial Sequence: TV1.8_5
US-10-190-435-4

Query Match	33.9%	Score 40;	DB 12;	Length 869;
Best Local Similarity	54.5%;	Pred. No. 6.7e+02;		
Matches	6;	Conservative	2;	Mismatches 3;
				Indels 0;
				Gaps 0;

QY	6	TYFPNKGTQY	16
		: :	
Db	399	TYYPKNGTYKY	409

RESULT 31
US-09-815-915-8

Sequence 8, Application US/09815915
Patent No. US20020025931A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
TITLE OF INVENTION: PROTEIN KINASE MOLECULES AND USES THEREFOR
FILE REFERENCE: 38155-20006.00
CURRENT APPLICATION NUMBER: US/09/815,915
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,846
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 8
LENGTH: 1244
TYPE: PRT
ORGANISM: Homo sapiens
US-09-815-915-8

Query Match	33.9%	Score 40;	DB 9;	Length 1244;
Best Local Similarity	46.2%	Pred. No. 1e+03;		
Matches	6;	Conservative	4;	Mismatches 3;
				Indels 0;
				Gaps 0;

```
QY      9 PNKGTOQYTDQIE 21
          |::|||::|
Db      626 PTASEQYTDRLIE 638
```

RESULT 32
US-09-924-154-17
; Sequence 17, Application US/09924154
; Patent No. US20020127241A1

TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match	33.9%	Score 40;	DB 10;	Length 1501;
Best Local Similarity	31.6%	Pred. No. 1.2e+03;		
Matches	6;	Conservative	6;	Mismatches 7;
				Indels 0;
				Gaps 0;

Qy 1 GLNMHTYFPNKGTOQTDD 19
 |::|||::: ::|
Db 1378 GVNSHLYNNLSSEBKMEQ 1396

RESULT 33

US-10-156-761-10289
; Sequence 10289, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10289

LENGTH: 107

TYPE: PRT
ORGANISM: Streptomyces avermitilis

US-10-156-761-10289

Query Match 33.1% Score 39; DB 15; Length 107;
Best Local Similarity 58.3% Pred. No. 95;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 HTYPEPKGTQYTDQIE 16

DB 61 HTTCFPGKGTASY 72

RESULT 34

US-10-156-761-9227
; Sequence 9227, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9227

LENGTH: 251

TYPE: PRT
ORGANISM: Streptomyces avermitilis

US-10-156-761-9227

Query Match 33.1% Score 39; DB 15; Length 251;
Best Local Similarity 33.3% Pred. No. 2.4e+02;

Matches 6; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 4 HTYPEPKGTQYTDQIE 21

DB 194 LRTLVPKSKRKRYKDRVE 211

RESULT 35

US-09-905-290A-4

Sequence 4, Application US/09905290A
Publication No. US20030087327A1
GENERAL INFORMATION:

APPLICANT: Rice, John
APPLICANT: Lanning, Beth
APPLICANT: Crawford, John
APPLICANT: Nye, Gordon
TITLE OF INVENTION: METHODS FOR MEASURING CYSTEINE AND DETERMINING CYSTEINE SYNTHASE
FILE REFERENCE: Docket No. US20030087327A12090US
CURRENT APPLICATION NUMBER: US/09/905,290A
CURRENT FILING DATE: 2001-07-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1

SEQ ID NO 4
LENGTH: 324

TYPE: PRT
ORGANISM: Arabidopsis thaliana

US-09-905-290A-4

Query Match 33.1% Score 39; DB 11; Length 324;
Best Local Similarity 38.1% Pred. No. 3.2e+02;

Matches 8; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GIANHTYPEPKGTQYTDQIE 21

DB 116 GAEVHLTPSKGKVGQIIDKAE 136

RESULT 36

US-09-738-626-6372
; Sequence 6372, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 6372

LENGTH: 380

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

US-09-738-626-6372

Query Match 33.1% Score 39; DB 10; Length 380;
Best Local Similarity 35.3% Pred. No. 3.9e+02;

Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 HTYPEPKGTQYTDQIE 21

DB 198 HSYGRNGEQAADBLQ 214

RESULT 37

US-09-738-626-4029

RESULT 39
US-10-116-519-8
; Sequence 8, Application US/10116519

Search completed: August 20, 2003, 10:08:16
Job time : 86.299 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 23.5979 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-2

Perfect score: 118

Sequence: 1 GNNMHTYFPNKGTOQYTDQIE 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	118	100.0	781	1 VCPV19	coat protein VP1 -
2	51	43.2	347	2 C90598	hypothetical prote
3	49.5	41.9	234	1 ZHBP79	head protein gp2 -
4	49.5	41.9	274	2 S04610	gene 2 protein - p
5	49	41.5	420	2 T32191	hypothetical prote
6	48	40.7	714	2 S76082	hypothetical prote
7	47	39.8	305	2 AB1321	probable phosphor
8	47	39.8	879	2 C83808	alanyl-tRNA synth
9	46	39.0	299	2 AG3278	hypothetical prote
10	46	39.0	458	2 B90625	NADH dehydrogenase
11	45	38.1	206	2 AB2152	transcription regu
12	45	38.1	206	2 C64634	iron-regulated oxi
13	45	38.1	611	2 D82881	zinc metalloprotei
14	45	38.1	789	2 T45762	vacuolar sorting p
15	45	38.1	815	2 F71880	probable iron-regu
16	44	37.3	315	2 T29258	hypothetical prote
17	44	37.3	324	2 E84972	hypothetical prote
18	44	37.3	357	2 T29771	hypothetical prote
19	44	37.3	445	2 AF3190	pyridoxal phosphat
20	44	37.3	617	1 S29742	malate dehydrogena
21	44	37.3	1002	2 S70292	FUN1 protein - ye
22	44	36.4	89	2 I68530	hemoglobin alpha c
23	43	36.4	185	2 S66113	hypothetical prote
24	43	36.4	256	2 S04363	class II histocomp
25	43	36.4	292	1 A39871	calponin alpha, sm
26	43	36.4	297	1 S31484	calponin HI - pig
27	43	36.4	297	1 S31486	calponin HI - mous
28	43	36.4	297	1 UN0773	calponin HI - rat
29	43	36.4	297	1 JC4500	basic calponin - h

30	43	36.4	297	2 G02142	smooth muscle cell
31	43	36.4	459	2 T11084	NADH2 dehydrogenas
32	43	36.4	459	2 T11333	NADH2 dehydrogenas
33	43	36.4	547	2 JC4519	heat-shock protein
34	43	36.4	548	2 C64076	chaperonin groEL -
35	43	36.4	564	2 A96999	pectate lyase rela
36	43	36.4	669	2 S14535	asparagine-rich pr
37	43	36.4	723	1 S30105	tryptophan 2-monoo
38	42.5	36.0	151	2 H70976	probable ribosomal
39	42.5	36.0	959	2 D83339	glycine cleavage s
40	42	35.6	88	2 A44929	protein-tyrosine-p
41	42	35.6	153	2 S72982	ribosomal protein
42	42	35.6	162	2 T17051	NADH2 dehydrogenas
43	42	35.6	170	2 T35564	ribosomal protein
44	42	35.6	238	2 S31779	trypsin (RC 3.4.21
45	42	35.6	283	2 E98061	conserved hypothet

ALIGNMENTS

RESULT 1

VCPV19 coat protein VP1 - parvovirus B19 (strain Au)

C:Species: parvovirus B19

A:Note: host Homo sapiens (man)

C:date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C:Accession: A24299

R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Astell, C.R.

J. Virol. 58, 921-936, 1986

A:title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr

A:Reference number: A24299; MUID:86200451; PMID:3701931

A:Accession: A24299

A:Molecule type: DNA

A:Residues: 1-781 <SHA>

A:Cross-references: EMBL:M13178; NID:G333375; PIDN:AAA66867.1; PID:G333377

C:Superfamily: parvovirus coat protein

C:Keywords: coat protein

Query Match	100.0%	Score 118;	DB 1;	Length 781;
Best Local Similarity	100.0%;	Pred. No. 3.2e-10;		
Matches	21;	Conservative	0;	Mismatches 0; Gaps 0;
Db	1			
	613	GNNMHTYFPNKGTOQYTDQIE	633	

RESULT 2

hypothetical protein MYPV 6910 (imported) - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: C90598

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.

Nucleic Acids Res. 29, 2145-2153, 2001

A:title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul

A:Reference number: A95512; MUID:21267165; PMID:11353084

A:Accession: C90598

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-347 <KUR>

A:Cross-references: GB:AL445566; PID:G14090106; PIDN:CAC13864.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 6910

A:Genetic code: SGC3

Query Match	43.2%	Score 51;	DB 2;	Length 347;
Best Local Similarity	47.4%;	Pred. No. 3.3;		
Matches	9;	Conservative	3;	Mismatches 7; Indels 0; Gaps 0;
Db	2			
	2	GNNMHTYFPNKGTOQYTDQI	20	

Db 304 LNNKHFFONRSTGTYIDQL 322

RESULT 3

ZHBPT9 head protein gp2 - phage T4

C:Species: phage T4

A:Note: host Escherichia coli

C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 23-Jul-1999

C/Accession: B32254

R/Author: B. Rao, A.S.M.K.; Bolten, B.M.; Balakrishnan, R.; Goldberg, E.B.

J. Bacteriol. 171, 488-497, 1989

A>Title: Cloning and identification of bacteriophage T4 gene 2 product gp2 and action of

A/Reference number: A91899; MUID:89123061; PMID:2644202

A/Accession: B32254

A/Molecule type: DNA

A/Residues: 1-234 <LIP>

A/Cross-references: GB:M23012; NID:9340786; PIDN:AAA50418.1; PID:9557582

C:Genetics:

C/Superfamily: phage T4 head protein gp2

C/Keywords: head protein

Query Match

41.9%; Score 49.5; DB 1; Length 234;

Best Local Similarity 45.0%; Pred. No. 3.8;

Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 1 GLNMHTYFPNKGTYQYTDQI 20

Db 136 GLNLMH-YIPPKARQFOEEL 154

RESULT 4

gene 2 protein - phage T4

C/Species: phage T4

C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Sep-1999

C/Accession: S04610

R/Koch, T.; Lamm, N.; Rueger, W.

Nucleic Acids Res. 17, 4392, 1989

A>Title: Sequencing, cloning and overexpression of genes of bacteriophage T4 between map

A/Reference number: S04608; MUID:89296504; PMID:2740234

A/Accession: S04610

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-274 <KOC>

A/Cross-references: EMBL:X14845; NID:915218; PIDN:CAA32951.1; PID:915221

C:Genetics:

C/Superfamily: phage T4 head protein gp2

Query Match

41.9%; Score 49.5; DB 2; Length 274;

Best Local Similarity 45.0%; Pred. No. 4.4;

Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 1 GLNMHTYFPNKGTYQYTDQI 20

Db 176 GLNLMH-YIPPKARQFOEEL 194

RESULT 5

T32191

hypothetical protein T01G6.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T32191

R/Jones, K.; Bradshaw, H.; Graves, T.

submitted to the EMBL Data Library, September 1997

A>Description: The sequence of C. elegans cosmid T01G6.

A/Reference number: Z21134

A/Accession: T32191

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-420 <JON>

A/Cross-references: EMBL:AF022978; PIDN:AA69073.1; GSPDB:GN00023; CESP:T01G6.6

A/Experimental source: strain Bristol N2; clone T01G6

C/Genetics:

A/Map position: 5

A/Introns: 61/2; 98/1; 234/3; 298/2; 387/3

Query Match

41.5%; Score 49; DB 2; Length 420;

Best Local Similarity 33.3%; Pred. No. 8.4;

Matches 6; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 3 NMHTYFPNKGTYQYTDQI 20

Db 359 DLHRYQDKGINRYERV 376

RESULT 6

S76082

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C/Species: Synechocystis sp.

A/Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 02-Aug-2002

C/Accession: S76082

R/Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud

DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

8.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S76082

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-714 <KAN>

A/Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BA10060.1; PID:d101071

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Genetics:

A/Start codon: GTG

C/Superfamily: Nostoc sp. hypothetical protein al12707; dnaI amino-terminal homology

F:6-70/Domain: dnaI amino-terminal homology <DNJ>

Query Match

40.7%; Score 48; DB 2; Length 714;

Best Local Similarity 45.0%; Pred. No. 21;

Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 LMHTYFPNKGTYQYTDQIE 21

Db 378 LSPKAYFNNPSVQYIEOLE 397

RESULT 7

AB1321

probable phosphotriesterase related proteins homolog lmo1970 [imported] - Listeria monoc

C/Species: Listeria monocytogenes

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C/Accession: AB1321

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Bagner, F.; Berche, P.; Bloeker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H

D.; Jones, L.M.; Karsel, U.

Science 294, 849-852, 2001

A/Authors: Krefel, T.; Kuhn, M.; Kunat, F.; Kurapat, G.; Madueno, E.; Maitouram, A.; Ma

ok, C.; Schueller, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A>Title: Comparative genomics of Listeria species

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AB1321

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-305 <GLA>

A/Cross-references: GB:NC_003210; PIDN:CAD00046.1; PID:g16411423; GSPDB:GN00177

A/Experimental source: strain EGD-e

C/Genetics:

A/Gene: lmo1970

A:Molecule type: DNA
A:Residues: 1-562 <ROM>
A:Cross-references: GB:AE000601; GB:AE000511; NID:g2314051; PIDN:AAD07961.1; PID:g231405

Query Match 38.1%; Score 45; DB 2; Length 562;
Best Local Similarity 47.4%; Pred. No. 48;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQ 19
||:|||||
DB 510 GLSLNMFVFNVPNQOYTDQ 528

RESULT 13

D82881
zinc metalloproteinase oligoendopeptidase F U521 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C/Accession: D82881

Class, J.I.; Lefkowitz, B.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
Submitted to GenBank, February 2000
Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870
A:Accession: D82881
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-611 <GUA>
A:Cross-references: GB:AE002150; GB:AF222894; NID:g6899515; PIDN:AAF30934.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
A:Genetics:
A:Gene: pepF-2; U521
A:Genetic code: SGC3

Query Match 38.1%; Score 45; DB 2; Length 611;
Best Local Similarity 38.9%; Pred. No. 52;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTD 18
||:|||||
DB 390 GHSVHTYFPNQSQEVNNE 407

RESULT 14

T45762
vacuolar sorting protein 35 homolog - Arabidopsis thaliana
N:Alternate names: protein F24M12.350
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 18-Feb-2000
C/Accession: T45762

Altale, D.; Litouri, R.; Flores, M.; Argitrou, A.; De Simone, V.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23012

A:Accession: T45762
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-789 <VIT>
A:Cross-references: EMBL:AL132980
A:Experimental source: cultivar Columbia; BAC clone F24M12
C:Genetics:

A:Map position: 3
A:Introns: 28/3; 76/3; 106/2; 168/2; 197/3; 216/3; 246/3; 287/3; 332/3; 371/3; 451/3; 47
A>Note: F24M12.350
C:Superfamily: Schistosaccharomyces probable vacuolar protein sorting-associated protein

Query Match 38.1%; Score 45; DB 2; Length 789;
Best Local Similarity 56.2%; Pred. No. 63;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNMHTYFPNKGTOOYTD 17
||:|||||
DB 717 LNKIYLFLEKGNQOYTD 732

RESULT 15
F71880
probable iron-regulated outer membrane protein - Helicobacter pylori (strain J99)
C/Species: Helicobacter pylori

A:Variety: strain J99
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
C/Accession: F71880
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Miller, S.D.; Jiang, Q.; Taylor, D.E.; Voyts, G.F.;
Nature 397, 176-180, 1999
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: F71880
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-815 <ARN>
A:Cross-references: GB:AE001514; GB:AE001439; NID:g4155418; PIDN:AAD06429.1; PID:g415542
A:Experimental source: strain J99
C:Genetics:
A:Gene: fmpB 2
C:Superfamily: iron-regulated outer membrane protein

Query Match 38.1%; Score 45; DB 2; Length 815;
Best Local Similarity 47.4%; Pred. No. 71;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQ 19
||:|||||
DB 763 GLSLNMFVFNVPNQOYTDQ 781

RESULT 16

T29258
hypothetical protein C01G8.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C/Accession: T29258

R:Du, Z.; Gattung, S.
submitted to the EMBL Data Library, November 1996
A:Description: The sequence of C. elegans cosmid C01G8.

A:Reference number: Z20597
A:Accession: T29258
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-315 <DUZ>
A:Cross-references: EMBL:U80439; PIDN:AA837638.1; GSPDB:GN00019; CESP:C01G8.1
A:Experimental source: strain Bristol N2; clone C01G8
C:Genetics:
A:Gene: CESP:C01G8.1

A:Map position: 1
A:Introns: 12/3; 66/1; 106/1; 170/3; 236/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C01G8.1

Query Match 37.3%; Score 44; DB 2; Length 315;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GLNMHTYFPN 10
||:|||||
DB 10 GLSTHAYFPN 19

RESULT 17

E84972
hypothetical protein [imported] - Buchnera sp. (strain APS)
C/Species: Buchnera sp.
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: E84972

R:Shigenobu, S.; Watanabe, H.; Hatori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A>Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. ;
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: E84972

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1324 <STO>
A:Cross-references: GB:AP000399; GSPDB:GN001444
A:Experimental source: strain APS
C:Genetics:
A:Gene: ycea, BU365
C:Superfamily: Escherichia coli ycea protein

Query Match	37.3%	Score 44	DB 2	Length 324
Best Local Similarity	42.1%	Pred. No. 38		
Matches	8	Conservative	5	Mismatches 6
				Indels 0
				Gaps 0

```
QY 2 LNMHTYFPNKGTOQYTDQI 20
    |::|||::|::
Db 27 LSFYKYFFIKNTQGYRDL 45
```

RESULT 18

Abstract: The *Caenorhabditis elegans* ZC581.1 protein is a member of the Cys-loop superfamily of ligand-gated ion channels. The ZC581.1 protein is expressed in the nervous system of *C. elegans* and is a functional ion channel. The ZC581.1 protein is a member of the Cys-loop superfamily of ligand-gated ion channels. The ZC581.1 protein is expressed in the nervous system of *C. elegans* and is a functional ion channel.

Query Match	37.3%	Score 44	DB 2	Length 357
Best Local Similarity	50.0%	Pred. 42		
Matches	9	Conservative	0	Mismatches 9
				Indels 0
				Gaps 0

QY 3 NMHTYFPNKGTOQYTDQI 20
|||
Db 104 NMREYFPPEKTVLDYFTQI 121

DLT 19

Pyridoxal phosphate aminotransferase [imported] - Agrobacterium tumefaciens (strain C58,
C:Accession: AF3190 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
S:Karp, P.; Romero, P.; Zhang, S.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
C:Accession: AF3190 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

Query Match	37.3%	Score 44;	DB 2;	Length 445;
Best Local Similarity	36.8%;	Pred. No. 53;		
Matches	7;	Conservative	6;	Mismatches 6; Indels 0; Gaps 0;

```
QY      2 LNMHTYFPNKGTOQYTDOI 20  
        ||:|| : | : ||| :  
Db      95 INIHTRYLNRVVEAYTENTL 113
```

RESULT 20

malate dehydrogenase (oxaloacetate-decarboxylating) (NADP) (EC 1.1.1.40) - pig roundworm
 N:Alternate names: malic enzyme
 C:Species: Ascaris suum (pig roundworm)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S29742
 R:Kulkarni, G.; Cook, P. F.; Harris, B. G.
 Arch. Biochem. Biophys. 300, 231-237, 1993
 A:Title: Cloning and nucleotide sequence of a full-length cDNA encoding Ascaris suum malate
 :Reference number: S29742; MIMD:53143319; PMID:8424657

Query Match	37.3%	Score 44;	DB 1;	Length 617;
Best Local Similarity	53.3%	Pred. No. 76;		
Matches	8;	Conservative	1;	Mismatches 6;
				Indels 0;
				Gaps 0;

QY 2 LNMHTYFPNKGTOY 16
||| | : |
Db 479 LNGTHYKPGQGNAY 493

RESULT 21

FUN12 protein (yeast (*Saccharomyces cerevisiae*))
 N:Alternate names: proteine YAL035w
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 02-Feb-2001
 C:Accession: S70292; S51984; S48519
 R:Vo, D.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: S70291
 A:Accession: S70292
 A:Molecule type: DNA
 A:Residues: 1-1002 <VOID>
 A:Cross-references: EMBL:U12980; NID:g1326053; PIDM:ANC04996.1; PID:g1326056; MIPS:YAL03.1
 R:Busey, H., Kaback, D.B., Zhong, W., Vo, D.T., Clark, M.W., Fortin, N., Hall, J., Quel
 submitted to the EMBL Data Library, August 1994
 .Description: The sequence of chromosome 1 of *Saccharomyces cerevisiae*.

A:Cross-references: SGD:S0000033; MIPS:YAL035W

A:Map position: 1L

C:Superfamily: translation elongation factor Tu homology

C:Keyword: GTP binding; nucleotide binding; P-loop; transmembrane protein

F:406-533/Domain: translation elongation factor Tu homology <ETU>

F:512-419/Region: nucleotide-binding motif A (P-loop)

F:491-507/Domain: transmembrane helix predicted <TM1>

F:530-533/Region: GTP-binding NKXD motif

Query Match 37.3%; Score 44; DB 2; Length 1002;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GUNMTYFPNKGTOQY 16

DB 576 GUNSELYFQNKMSKY 591

RESULT 22

530

C:Superfamily: globin alpha chain (clone alpha28) - rat (fragment)

C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 19-May-2000

C:Accession: I68530

R:Ham, V.M.; Gu, Y.L.; Au, D.M.; Wong, W.M.; Ma, C.W.; Cheng, L.Y.

Hemoglobin 17, 363-371, 1993

A:Title: Two new rat alpha-globin sequences as identified by the conserved region PCR.

A:Reference number: I54239; MUID:94042225; PMID:8226096

A:Accession: I68530

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-89 <RES>

A:Cross-references: GB:S66558; NID:9439804

C:Genetics: 32/2

A:Introns: 32/2

C:Superfamily: globin; globin homology

C:Keyword: blood; oxygen carrier

F:59/Binding site: oxygen (His) (distal axial ligand) #status predicted

Query Match 36.4%; Score 43; DB 2; Length 89;

Best Local Similarity 43.8%; Pred. No. 14;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 6 TYPNKGTOQYTOIE 21

DB 42 TYPHNTSGSDQV 57

DB 42 TYPHNTSGSDQV 57

RESULT 23

513

C:Superfamily: hypothetical protein yach - Bacillus subtilis

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C:Accession: S66113; A69741

R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.

DNA Res. 1, 1-14, 1994

A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom

A:Reference number: S65867; MUID:96051385; PMID:7584024

A:Accession: S66113

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-185 <COA>

A:Cross-references: EMBL:026185; NID:9467326; PIDN:BA05318.1; PID:9467472

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993

R:Kunze, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beretz

C:Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A:Enllich, S.D.; Emerson, P.T.; Enlian, K.D.; Errington, J.; Fabref, C.; Ferrari, B.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier

koeth, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hulo, M.F.

Koetter, P.; Konigstein, G.; Ktogh, S.; Kumano, M.; Kurita, A.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Mauda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelid

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror

akouchi, M.; Tanakehi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumelein, B.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: A69741

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-185 <KUN>

A:Cross-references: GB:Z99104; GB:AL009126; NID:92632267; PIDN:CA01860.1; PID:92632351

A:Experimental source: strain 168

C:Genetics: yach

A:Gene: yach

A:Start codon: TTG

C:Superfamily: Bacillus subtilis hypothetical protein yach

Query Match 36.4%; Score 43; DB 2; Length 185;

Best Local Similarity 50.0%; Pred. No. 30;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 LNMHTYFPNKGTOQYT 17

DB 58 LNMDSFQNAQTQWFS 73

RESULT 24

504363

C:Superfamily: class II histocompatibility antigen RT1-B alpha chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000

C:Accession: S04363

R:Syha, J.; Henkes, W.; Reske, K.

Nucleic Acids Res. 17, 3985, 1989

A:Title: Complete cDNA sequence coding for the MHC class II RT1.B alpha chain of the LEW

A:Reference number: S04363; MUID:89282410; PMID:2499874

A:Accession: S04363

A:Molecule type: mRNA

A:Residues: 1-256 <SYH>

A:Cross-references: EMBL:X14879; NID:957154; PIDN:CAA33020.1; PID:957155

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-256/Product: class II histocompatibility antigen, RT1-B alpha chain #status predict

F:127-192/Domain: immunoglobulin homology <IMM>

Query Match 36.4%; Score 43; DB 2; Length 256;

Best Local Similarity 33.3%; Pred. No. 42;

Matches 7; Conservative 8; Mismatches 4; Indels 2; Gaps 1;

QY 1 GUNMTYFPNKGTOQYTOIE 21

DB 36 GUNMTQYBSRG--OFTHFPD 54

RESULT 25

A39871

C:Superfamily: calponin alpha, smooth muscle - chicken

N:Contamin: calponin beta, smooth muscle

C:Species: Gallus gallus (chicken)

C:Date: 30-Dec-1991 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999

C:Accession: A39871; B39871; A46007; A42859; S39778

R:Takashi, K.; Nadal-Ginard, B.

J. Biol. Chem. 266, 13284-13288, 1991

A:Title: Molecular cloning and sequence analysis of smooth muscle calponin.

A:Reference number: A39871; MUID:91302359; PMID:2071603

A:Accession: A39871

A:Molecule type: mRNA

A:Residues: 1-292 <TAK>

A:Cross-references: GB:M63559; NID:9211387; PIDN:AAA48651.1; PID:9211388

A:Accession: B39871

A:Molecule type: mRNA

A:Residues: 1-216, 257-292 <TA2>

A:Cross-references: GB:M63560; NID:g211389; PIDN:AAA48652.1; PID:g211390
 R.Nakamura, F.; Mino, T.; Yamamoto, J.; Naka, M.; Tanaka, T.
 J. Biol. Chem. 268, 6194-6201, 1993
 A:Title: Identification of the regulatory site in smooth muscle calponin that is phospho-
 A:Reference number: A46007; MUID:93203204; PMID:8454594
 A:Accession: A46007
 A:Molecule type: protein
 A:Residues: 173-185;213-225;252-265 <NAK>
 R.Mesquedel, M.; Patoum, A.; Derancourt, J.; Kaesab, R.
 J. Biol. Chem. 267, 15943-15951, 1992
 A:Title: Mapping of the functional domains in the amino-terminal region of calponin.
 A:Reference number: A42859; MUID:92348464; PMID:1639822
 A:Accession: A42859
 A:Molecule type: protein
 A:Residues: 7-21;52-65; 'K';183-194 <ME>
 A:Note: authors found Thr-184 to be the preferred and functionally most important site for
 R.Winder, S.J.; Allen, B.G.; Frazer, E.D.; Kang, H.M.; Kargacin, G.J.; Walsh, M.P.
 Biochem. J. 296, 827-836, 1993
 A:Title: Calponin phosphorylation in vitro and in intact muscle.
 A:Reference number: S39778; MUID:94107251; PMID:8280082
 A:Accession: S39778
 A:Molecule type: protein
 A:Residues: 67-74;173-185;252-256 <MIN>
 A:Note: authors found phosphorylation at Ser-175, the major site for modification by pro-
 C:Comment: This thin filament-associated protein binds actin, tropomyosin, and calmodulin
 myosin. Phosphorylation of calponin causes dissociation from actin and tropomyosin and
 C:Superfamily: calponin; calponin repeat homology; smooth muscle protein SM22 homology
 C:Keywords: actin binding; alternative splicing; calmodulin binding; muscle contraction;
 F:1-392/Domain: actin binding; smooth muscle #status predicted <CALM>
 F:1-216;257-292/Domain: calponin beta, smooth muscle #status predicted <CALM>
 F:6-190/Domain: smooth muscle protein SM22 homology <SMH>
 F:52-144/Region: calmodulin binding #status predicted
 F:145-182/Region: actin binding #status predicted
 F:164-194/Domain: calponin repeat homology <CRH1>
 F:204-234/Domain: calponin repeat homology <CRH2>
 F:243-273/Domain: calponin repeat homology <CRH3>

Query Match 36.4%; Score 43; DB 1; Length 292;
 Best Local Similarity 47.4%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GNNMHTYPPNKGTOGYTDQ 19
 Db 183 GTRRLYDPKLGTDQPLDQ 201

RESULT 26
 A:Species: Sus scrofa domestica (domestic pig)
 C:Date: 20-Feb-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999
 C:Accession: S36145; S31484
 R:Strasser, P.; Glimona, M.; Moessler, H.; Herzog, M.; Small, J.V.
 FEBS Lett. 330, 13-18, 1993
 A:Title: Mammalian calponin. Identification and expression of genetic variants.
 A:Reference number: S36145; MUID:93380560; PMID:8370452
 A:Accession: S36145
 A:Molecule type: mRNA
 A:Residues: 1-297 <ST>
 A:Cross-references: EMBL:Z19538; NID:g1959; PIDN:CAA79598.1; PID:g1960
 C:Comment: This thin filament-associated protein binds actin, tropomyosin, and calmodulin
 myosin. Phosphorylation of calponin causes dissociation from actin and tropomyosin and
 C:Superfamily: calponin; calponin repeat homology; smooth muscle protein SM22 homology
 C:Keywords: actin binding; calmodulin binding; muscle contraction; phosphoprotein; smooth
 F:6-190/Domain: smooth muscle protein SM22 homology <SMH>
 F:164-194/Domain: calponin repeat homology <CRH1>
 F:204-234/Domain: calponin repeat homology <CRH2>
 F:243-273/Domain: calponin repeat homology <CRH3>

Query Match 36.4%; Score 43; DB 1; Length 297;
 Best Local Similarity 47.4%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GNNMHTYPPNKGTOGYTDQ 19
 Db 183 GTRRLYDPKLGTDQPLDQ 201

RESULT 27
 A:Species: Mus musculus (house mouse)
 C:Date: 20-Feb-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999
 C:Accession: S36145; S31486
 R:Strasser, P.; Glimona, M.; Moessler, H.; Herzog, M.; Small, J.V.
 FEBS Lett. 330, 13-18, 1993
 A:Title: Mammalian calponin. Identification and expression of genetic variants.
 A:Reference number: S36145; MUID:93380560; PMID:8370452
 A:Accession: S36145
 A:Molecule type: mRNA
 A:Residues: 1-297 <ST>
 A:Cross-references: EMBL:Z19542; NID:g51137; PIDN:CAA79602.1; PID:g51138
 C:Comment: This thin filament-associated protein binds actin, tropomyosin, and calmodulin
 myosin. Phosphorylation of calponin causes dissociation from actin and tropomyosin and
 C:Superfamily: calponin; calponin repeat homology; smooth muscle protein SM22 homology
 C:Keywords: actin binding; calmodulin binding; muscle contraction; phosphoprotein; smooth
 F:6-190/Domain: smooth muscle protein SM22 homology <SMH>
 F:164-194/Domain: calponin repeat homology <CRH1>
 F:204-234/Domain: calponin repeat homology <CRH2>
 F:243-273/Domain: calponin repeat homology <CRH3>

Query Match 36.4%; Score 43; DB 1; Length 297;
 Best Local Similarity 47.4%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GNNMHTYPPNKGTOGYTDQ 19
 Db 183 GTRRLYDPKLGTDQPLDQ 201

RESULT 28
 A:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1993 #sequence_revision 07-Jul-1995 #text_change 16-Jun-2000
 C:Accession: S37640; JN0773
 R:Shanahan, C.M.; Weisberg, P.L.; Metcalfe, J.C.
 Circ. Res. 73, 193-204, 1993
 A:Title: Isolation of gene markers of differentiated and proliferating vascular smooth m
 A:Reference number: S37637; MUID:93284726; PMID:8508530
 A:Accession: S37640
 A:Molecule type: mRNA
 A:Residues: 1-297 <SHA>
 A:Cross-references: EMBL:X71071; NID:g313817; PIDN:CAA50397.1; PID:g313818
 R:Nishida, W.; Kitami, Y.; Hwada, K.
 Gene 130, 287-302, 1993
 A:Title: cDNA cloning and mRNA expression of calponin and SM22 in rat aorta smooth muscle
 A:Reference number: JN0773; MUID:93366190; PMID:8359698
 A:Accession: JN0773
 A:Molecule type: mRNA
 A:Residues: 1-96; 'E'; 287-297 <NIS>
 A:Cross-references: GB:D14437; NID:g436047; PIDN:BA403320.1; PID:g436048
 C:Comment: This thin filament-associated protein binds actin, tropomyosin, and calmodulin
 myosin. Phosphorylation of calponin causes dissociation from actin and tropomyosin and
 C:Superfamily: calponin; calponin repeat homology; smooth muscle protein SM22 homology
 C:Keywords: actin binding; calmodulin binding; muscle contraction; phosphoprotein; smooth
 F:6-190/Domain: smooth muscle protein SM22 homology <SMH>
 F:145-182/Region: calmodulin binding #status predicted
 F:145-182/Region: actin binding #status predicted
 F:164-194/Domain: calponin repeat homology <CRH1>
 F:204-234/Domain: calponin repeat homology <CRH2>
 F:243-273/Domain: calponin repeat homology <CRH3>

Query Match 36.4%; Score 43; DB 1; Length 297;
 Best Local Similarity 47.4%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQ 19
 DB 183 GTRRHLYDPKLGTDQPLDQ 201

RESULT 29

JC4500
 basic calponin - human
 C:Species: Homo sapiens (man)
 C>Date: 07-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 28-May-1999
 C:Accession: J04500
 R:Magnuhl, M.; Nishida, W.; Kohara, K.; Kuwano, A.; Kondo, I.; Hiwada, K.
 Biochem. Biophys. Res. Commun. 217, 238-244, 1995
 A:Title: Molecular cloning and gene mapping of human basic and acidic calponins.
 A:Reference number: J04500; MUID:96095663; PMID:8526917
 A:Accession: J04500
 A:Molecule type: mRNA
 A:Residues: 1-297 <MAG>
 A:Cross-references: GB:S80560; NID:g1245964; PIDN:AAB35751.1; PID:g1245965
 A:Experimental source: Aorta
 A:Genetics:

A:Map position: 19p13.1-13.2
 C:Superfamily: calponin; calponin repeat homology; smooth muscle protein SM22 homology
 C:Keywords: actin binding; tandem repeat
 F:5-190/Domain: smooth muscle protein SM22 homology <SMH>
 F:164-194/Domain: calponin repeat homology <CRH1>
 F:204-234/Domain: calponin repeat homology <CRH2>
 F:243-273/Domain: calponin repeat homology <CRH3>

Query Match 36.4%; Score 43; DB 2; Length 297;
 Best Local Similarity 47.4%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQ 19
 DB 183 GTRRHLYDPKLGTDQPLDQ 201

RESULT 30

G02142
 smooth muscle cell calponin - human
 C:Species: Homo sapiens (man)
 C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 21-Jan-2000
 C:Accession: G02142
 R:Miano, J.N.; Krahe, R.; Olson, E.N.
 submitted to the EMBL Data Library, September 1995
 A:Reference number: G09218
 A:Accession: G02142
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-297 <MIA>
 A:Cross-references: EMBL:U37019; NID:g1017775; PIDN:AAC51780.1; PID:g1017776
 C:Superfamily: calponin; calponin repeat homology; smooth muscle protein SM22 homology
 F:5-190/Domain: smooth muscle protein SM22 homology <SMH>
 F:164-194/Domain: calponin repeat homology <CRH1>

Query Match 36.4%; Score 43; DB 2; Length 297;
 Best Local Similarity 47.4%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQ 19
 DB 183 GTRRHLYDPKLGTDQPLDQ 201

RESULT 31

T11084
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Falco peregrinus mitochondrion
 C:Species: mitochondrion Falco peregrinus
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
 C:Accession: T11084
 R:Windell, D.P.; Sorenson, M.D.; Dimcheff, D.E.

Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
 A:Title: Multiple independent origins of mitochondrial gene order in birds.
 A:Reference number: Z17242

A:Accession: T11084
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-459 <MIN>
 A:Cross-references: EMBL:AF090338; NID:g4894462; PID:g4894472; PIDN:AAD32501.1
 C:Genetics:

A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 36.4%; Score 43; DB 2; Length 459;
 Best Local Similarity 44.0%; Pred. No. 79;
 Matches 11; Conservative 5; Mismatches 3; Indels 6; Gaps 2;

QY 2 LNMTYFPNKGTOQYTDQ 20
 DB 37 LSLHWLPLTPYPSKTLTQMTGMDQ 61

RESULT 32

T11333
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Corvus frugilegus mitochondrion
 C:Species: mitochondrion Corvus frugilegus
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
 C:Accession: T11333
 R:Harlid, A.; Arnason, U.
 Proc. R. Soc. Lond. B Biol. Sci. 266, 305-309, 1999
 A:Title: Analyses of mitochondrial DNA nest rattle birds within the Neognathae-supportin
 A:Reference number: Z17262

A:Accession: T11333
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-459 <HAR>
 A:Cross-references: EMBL:Y18522; PIDN:CAV7204.1
 C:Genetics:

A:Genome: mitochondrion
 A:Note: NADH4
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 36.4%; Score 43; DB 2; Length 459;
 Best Local Similarity 52.9%; Pred. No. 79;
 Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 6 TYFPNKGTOQYTDQ 20
 DB 45 TYPSKGLTHWTSIDQ 61

RESULT 33

JC4519
 heat-shock protein GroEL - Pasteurella multocida
 C:Species: Pasteurella multocida
 C>Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
 C:Accession: J04519
 R:Love, B.C.; Hansen, L.M.; Hirsh, D.C.
 Gene 166, 179-180, 1995
 A:Title: Cloning and sequence of the groEL heat-shock operon of Pasteurella multocida.
 A:Reference number: J04518; MUID:96105224; PMID:8529887
 A:Accession: J04519

A:Molecule type: DNA
 A:Residues: 1-547 <LOV>
 A:Cross-references: GB:U30165; NID:g1144300; PIDN:AAA84916.1; PID:g1144302
 A:Experimental source: serotype A:3
 C:Genetics:

A:Gene: groEL
 C:Superfamily: chaperonin groEL
 C:Keywords: heat shock; stress-induced protein

C/Genetics:
A/Gene: rpsI

C:Superfamily: Escherichia coli ribosomal protein S9

Job time : 26.5979 secs

Query Match	36.0%	Score 42.5;	DB 2;	Length 151;
Best Local Similarity	35.7%	Pred. No. 29;		
Matches 10;	Conservative 4;	Mismatches 7;	Indels 7;	Gaps 1.

```

QY      1 GLNHTTYFPNKGTOQY-----TDQIE 21
          |::||| | |
Db      53 GRSLEDYFPNKKVHQQLIKAPLVTVDRVE 80

```

RESULT 39
D03330

glycine cleavage system protein P2 PA2445 [imported] - Pseudomonas aeruginosa (strain PR4)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Dec-2000#cont shared 31-Dec-2000

C:\Date: 15-Sep-2000 #sequence_revision: 15-Sep-2000 #text_change 31-Dec-2000
C:\Accession: D83339

R. Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warener, P.; Hickey, M.J.; Baseman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

JOSEPH, S.; OLSON, M.V.
Box 406, 959-964, 2000
141-142-143-144-145-146-147-148-149-150-151-152-153-154-155-156-157-158-159-160-161-162-163-164-165-166-167-168-169-170-171-172-173-174-175-176-177-178-179-180-181-182-183-184-185-186-187-188-189-190-191-192-193-194-195-196-197-198-199-200-201-202-203-204-205-206-207-208-209-210-211-212-213-214-215-216-217-218-219-220-221-222-223-224-225-226-227-228-229-230-231-232-233-234-235-236-237-238-239-240-241-242-243-244-245-246-247-248-249-250-251-252-253-254-255-256-257-258-259-260-261-262-263-264-265-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000-1001-1002-1003-1004-1005-1006-1007-1008-1009-1010-1011-1012-1013-1014-1015-1016-1017-1018-1019-1020-1021-1022-1023-1024-1025-1026-1027-1028-1029-1030-1031-1032-1033-1034-1035-1036-1037-1038-1039-1040-1041-1042-1043-1044-1045-1046-1047-1048-1049-1050-1051-1052-1053-1054-1055-1056-1057-1058-1059-1060-1061-1062-1063-1064-1065-1066-1067-1068-1069-1070-1071-1072-1073-1074-1075-1076-1077-1078-1079-1080-1081-1082-1083-1084-1085-1086-1087-1088-1089-1090-1091-1092-1093-1094-1095-1096-1097-1098-1099-1100-1101-1102-1103-1104-1105-1106-1107-1108-1109-1110-1111-1112-1113-1114-1115-1116-1117-1118-1119-1120-1121-1122-1123-11

Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: NC0110

A;Accession: D833339
A;Status: preliminary
A;Molecular type: DNA

A;Molecule type: DNA
A;Residues: 1-959 <STO>

A:Cross-references: GB:AE004672; GB:AE004091; NID:g9948422; PIDN:AAG5833.1; GSPDE:SN001
A:Experimental source: strain PA01
C:Genetics:

C;Genetic8:
A;Gene: gcvp2; PA2445

Query Match	36.0%;	Score 42.5;	DB 2;	Length 959;
Best Local Similarity	36.4%;	Prod No 3 10-03;		

Matches 8; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

```

QY      3 NMHTYFP---NKGTQYTDQIE 21
          |::|::|::|::|
Db      535 NLHPAPAEQSEGRYQLTDELE 556

```

RESULT 40

protein-tyrosine-phosphatase (EC 3.1.3.48) BPTP-1 - human (fragment)

C:\Species: Homo sapiens (man)
C:\Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Mar-1998

C:Accession: A44929
M: Sekiya, M.; Arimura, Y.; Takekawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yano, T.
F: 1997
P: 1000

Cell Res. 52, 737-740, 1992
Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6
Number: 74000
Date: 01/05/92
Page: 17001

A;Reference number:
A;Accession: A44929

A;Molecule type: mRNA
A;Residues: 1-88 <ADA>

A;Experimental source: pre-B cell NALM-6

A;Note: the authors translated the codon GGC for residue 12 as Glu
A;Note: the authors did not report the entire codon for residue 88
A;Note: the authors did not report the entire codon for residue 90

C;superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology#
oy
C;superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology#
oy

Keywords: phosphoric monoester hydrolase; tyrosine-specific phosphatase C1-88/Domain: leukocyte common antigen cytosolic domain homology (fragment) <LAC-

Query Match	35.6%	Score 42;	DB 2;	Length 88;
Best Local Similarity	60.0%	Score 10;		

Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0.

QY	7	YFPNKGTOQY	16
		: : : :	
Db	2	YMPNRGTETY	11

Search completed: August 20, 2003, 09:32:34

DE Terminal DNA protecting protein (Head protein Gp2) (Gp64).
 GN 2 OR 64.
 OS Bacteriophage T4.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses
 ON NCBI_TaxId=10665;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89296504; PubMed=2740234;
 RA Koch T., Lamm N., Rueger W.;
 RT "Sequencing, cloning and overexpression of genes of bacteriophage T4
 RL Nucleic Acids Res. 17:4392-4392(1989)."
 RN
 RP SEQUENCE FROM N.A.
 RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
 RA Mesyanzhinov V., Ruger W., Strickham T., Thomas E.;
 RT "Bacteriophage T4 genome analysis."
 RN Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN
 [3]
 CC SEQUENCE OF 41-234 FROM N.A.
 CC MEDLINE=89123061; PubMed=2644202;
 CC Lipinska B., Rao A.S.M.K., Bolten B.M., Balakrishnan R.,
 CC Goldberg E.B.;
 CC "Cloning and identification of bacteriophage T4 gene 2 product gp2
 CC and action of gp2 on infecting DNA in vivo."
 CC J. Bacteriol. 171:488-497(1989).
 CC
 CC -1- FUNCTION: THIS PROTEIN IS A HEAD PROTEIN OF BACTERIOPHAGE T4.
 CC IT PROTECTS T4 DNA AGAINST EXO MEDIATED DEGRADATION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X14845; CAA32951.1; -
 CC DR EMBL; AF158101; AAD2438.1; -
 CC DR EMBL; M23012; AAA50418.1; ALT_INIT.
 CC DR PIR; S04610; S04610.
 CC
 CC Structural protein.
 CC FT CONFLICT 71 AERIKQ -> VKRFKN (IN REF. 3).
 CC FT CONFLICT 118 R -> K (IN REF. 3).
 CC FT CONFLICT 191 L -> Q (IN REF. 3).
 CC SQ SEQUENCE 274 AA; 31610 MW; 3583A0A76439B3B1 CRC64;
 CC
 CC Query Match 41.9%; Score 49.5; DB 1; Length 274;
 CC Best Local Similarity 45.0%; Pred. No. 1.5;
 CC Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;
 CC
 CC 1 GLNMHTYFPNKGQOYTDQI 20
 CC Db 176 GLNMHTYFPNKGQOYTDQI 194
 CC
 CC RESULT 3
 CC SYA_BACHD STANDARD; PRT; 879 AA.
 CC AC 09KDE6;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Alanyl-tRNA synthetase (Ec 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
 CC GN ALAS OR BH1267.
 CC OS Bacillus halodurans.
 CC OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC ON NCBI_TaxId=86655;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN=C-125 / JCM 9153;
 CC RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Saeaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC
 CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
 CC diphosphate + L-alanyl-tRNA(Ala).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP001511; BAB04986.1; -
 CC DR PIR; C83808; C83808.
 CC DR HAMAP; MF_00036; -; 1.
 CC DR InterPro; IPR003156; DHAH1.
 CC DR InterPro; IPR002318; tRNA-synt_2c.
 CC DR InterPro; IPR006193; tRNA-synt_Ala.
 CC DR Pfam; PF02272; DHAH1; 1.
 CC DR Pfam; PF01411; tRNA-synt_2c; 1.
 CC DR PRINTS; PR00980; TRNASYNTHALA.
 CC DR TRIGRAMS; TRIGR00344; alas; 1.
 CC DR PROSITE; PS00860; AA_tRNA_LIGASE_II_ALA; 1.
 CC KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding;
 CC Complete proteome.
 CC SQ SEQUENCE 879 AA; 97741 MW; 4BC9C5CC397A7F41 CRC64;
 CC
 CC Query Match 39.8%; Score 47; DB 1; Length 879;
 CC Best Local Similarity 38.1%; Pred. No. 13;
 CC Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 CC
 CC 1 GLNMHTYFPNKGQOYTDQIE 21
 CC Db 531 GQHLHTVYKESGLQVNDQVQ 551
 CC
 CC RESULT 4
 CC Y365_BUCAI STANDARD; PRT; 324 AA.
 CC AC P57456;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Hypothetical protein BUI365.
 CC GN BUI365.
 CC OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 CC symbiotic bacterium).
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC OC Enterobacteriaceae; Buchnera.
 CC ON NCBI_TaxId=118099;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN=Tokyo 1998;
 CC RX MEDLINE=20445173; PubMed=10993077;
 CC RA Shigenobu S., Watanabe H., Hatford M., Sakaki Y., Ishikawa H.;
 CC "Genome sequence of the endocellular bacterial symbiont of aphids
 CC Buchnera sp. APS."
 CC RT Nature 407:81-86(2000).
 CC RL Nature 407:81-86(2000).
 CC
 CC -1- SIMILARITY: BELONGS TO THE UPF0176 FAMILY.
 CC -1- SIMILARITY: Contains 1 rhodanese domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

CC CC entties requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
DR DR EMBL; AP001119; BABI3069.1; -.
DR DR HAMAP; MF_00469; -; 1.
DR DR InterPro; IPR001763; Rhodanese-like.
DR DR Pfam; PF00581; Rhodanese; 1.
DR DR SMART; SM00450; RHOD; 1.
DR DR PROSITE; PS50206; RHODANESE_3; 1.
DR DR Hypothetical protein; Complete proteome.
KW DOMAIN 145 239 RHODANESE.
SQ SEQUENCE 324 AA; 38405 MW; 9E7B019857D377FA CRC64;

Query Match 37.3%; Score 44; DB 1; Length 324;
Best Local Similarity 42.1%; Pred. No. 13;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

2 LNMHTYFPNKGTOQYTDPI 20
27 LSPYKTFPIKNTQETRDRL 45

RESULT 5
NAOM_ASCSU STANDARD; PRT; 643 AA.
ID NAOM_ASCSU
AC P27443;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NAD-dependent malic enzyme, mitochondrial precursor (EC 1.1.1.38)
DE (NAD-ME) (Fragment).
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoidea).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_Taxid=6253;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93143319; PubMed=8424657;
RA Kulikarni G., Cook P.F., Harris B.G.;
RT "Cloning and nucleotide sequence of a full-length cDNA encoding
RT Ascaris suum malic enzyme.";
RL Arch. Biochem. Biophys. 300:231-237(1993).
CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = pyruvate + CO(2) + NADH.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: THIS ISOENZYME CAN ALSO USE NADP(+) BUT IS MORE
CC EFFECTIVE WITH NAD(+).
CC -1- SIMILARITY: BELONGS TO THE MALIC ENZYMES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC CC EMBL; M81055; -; NOT ANNOTATED_CDS.
DR PDB; 1LLO; 07-AUG-02_
DR InterPro; IPR001891; Malic_oxred.
DR DR Pfam; PF00390; malic_1.
DR DR Pfam; PF03949; malic_N; 1.
DR DR PRINTS; PR00072; MALOXRDPTASE.
DR DR PROSITE; PS00331; MALIC_ENZYMES; 1.
KW Oxidoreductase; NAD; Mitochondrion; Transit peptide; 3D-structure.
FT NON TER 1 1
FT TRANSIT <1 39 MITOCHONDRION (POTENTIAL).
FT CHAIN 40 643 NAD-DEPENDENT MALIC ENZYME.
SQ SEQUENCE 643 AA; 72757 MW; 7B1A480F086FE267 CRC64;

Query Match 37.3%; Score 44; DB 1; Length 643;
Best Local Similarity 53.3%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

```

Cy      2  LMHATFENKGTQCY 16
      || || || |
Db      505  LNGTHYFPGCGNNAY 519

RESULT 6
IF2P_YEAST
ID _IF2P_YEAST STANDARD; PRT; 1002 AA.
AC P39730.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Translation initiation factor IF-2.
GN YAL035W OR FUN12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=95249563; PubMed=7731988;
RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
RA Hall J., Quellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
RA Storms R.K.;
RT "The nucleotide sequence of chromosome I from Saccharomyces
RT cerevisiae."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=94357438; PubMed=8076820;
RA Surraive P., Shafer B.K., Strathern J.N., Hughes S.H.;
RT "Isolation, identification and characterization of the FUN12 gene of
RT Saccharomyces cerevisiae."
RL Gene 146:209-213(1994).
[3]
RN REVISIONS.
RC STRAIN=S288c / AB972;
RA Vo D.;
RN Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[4]
RN CHARACTERIZATION.
RX MEDLINE=98288357; PubMed=9624054;
RA Choi S.K., Lee J.H., Zoll W.L., Merrick W.C., Deyer T.E.;
RT "Promotion of met-rRNA:met binding to ribosomes by yIF2, a bacterial
RT IF2 homolog in yeast."
RL Science 280:1757-1760(1998).
-1- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING
THE BINDING OF THE FORMYL METHIONINE-TRNA TO RIBOSOMES. SEEMS TO
FUNCTION ALONG WITH EIF-2.
-1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U12980; AAC04966.1; -
DR EMBL; L29389; AAA57228.1; ALT_SEQ.
DR PIR; S70282; S70282.
DR SCD; S0000033; FUN12.
DR GO; GO:0005843; C:cytosolic small ribosomal subunit (sensu Eu. .; IDA.
DR GO; GO:0003743; F:translation initiation factor activity; IDA.
DR GO; GO:0006413; P:translation initiation; IDA.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000178; IF2.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.

```

DR Pfam; PF03144; GTP EFTU D2; 1.
 DR PRINTS; PR00315; ELONGINFCT.
 DR TRIGRAME; TRIGR00231; small GTP; 1.
 DR PROSITE; PS01176; IF2; FALSE NEG.
 KM Initiation factor; Protein biosynthesis; GTP-binding.
 FT DOMAIN 361 371
 FT NP BIND 412 419 GTP (BY SIMILARITY).
 SQ SEQUENCE 1002 AA; 112268 MW; 1A496195DAE1C283 CRC64;

Query Match 37.3%; Score 44; DB 1; Length 1002;
 Best Local Similarity 50.0%; Pred. No. 45;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GLNMHTYFPNKGTOQY 16
 Db 576 GLNSELVFNKNSKY 591

RESULT 7
 H_BACSU STANDARD; PRT; 185 AA.
 P37559;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yach.
 GN YACH.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96051385; PubMed=7584024;
 RA Ogasawara N., Nakai S., Yoshikawa H.;
 RT "Systematic sequencing of the 180 kilobase region of the Bacillus
 RL subtilis chromosome containing the replication origin";
 DNA Res. 1:1-14 (1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berteiro M.G., Bessieres P., Bolochin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Duesterhoft A., Enright S.D., Emerson P.T.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Katsanara Y., Klaer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche R., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin J., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Teipstra P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Welfzenegger T.,
 RA Winzars P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RL subtilis";
 RJ Nature 390:249-256 (1997).

CC -1- SIMILARITY: Contains 1 UVR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL; D26185; BAA05318.1; -
 CC EMBL; Z59104; CAB11860.1; -
 CC PIR; S66113; S66113.
 CC Subtilist; BG10146; Yach.
 CC InterPro; IPR001943; UVRB/C.
 CC DR Pfam; PF02151; UVR; 1.
 CC PROSITE; PS50151; UVR; 1.
 CC KW Hypothetical protein; Complete proteome.
 CC FT DOMAIN 139 174 UVR.
 CC SEQUENCE 185 AA; 21023 MW; A6388713A1A6ADB CRC64;

Query Match 36.4%; Score 43; DB 1; Length 185;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LNMHTYFPNKGTOQY 17
 Db 58 LNMDSFQNAQTQMS 73

RESULT 8
 CLPO CHICK STANDARD; PRT; 292 AA.
 AC P26932; P26933;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calponin alpha and beta, smooth muscle.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=gizzard smooth muscle;
 RX MEDLINE=91302359; PubMed=2071603;
 RA Takahashi K., Nadal-Ginard B.;
 RT "Molecular cloning and sequence analysis of smooth muscle calponin.";
 RL J. Biol. Chem. 268:13284-13286 (1991).
 [2]
 RP PARTIAL SEQUENCE, AND MUTAGENESIS.
 RX MEDLINE=94103161; PubMed=8276752;
 RA Gong B.J., Mabuchi K., Takahashi K., Nadal-Ginard B., Tao T.;
 RT "Characterization of wild type and mutant chicken gizzard alpha
 RL calponin expressed in E. coli.";
 RL J. Biochem. 114:453-456 (1993).
 [3]
 RP PHOSPHORYLATION.
 RX MEDLINE=93203204; PubMed=8454594;
 RA Nakamura F., Mino T., Yamamoto J., Naka M., Tanaka T.;
 RT "Identification of the regulatory site in smooth muscle calponin that
 RL is phosphorylated by protein kinase C.";
 RL J. Biol. Chem. 268:6194-6201 (1993).
 CC -1- FUNCTION: THIN FILAMENT-ASSOCIATED PROTEIN THAT IS IMPLICATED IN
 CC THE REGULATION AND MODULATION OF SMOOTH MUSCLE CONTRACTION. IT IS
 CC CAPABLE OF BINDING TO ACTIN, CALMODULIN, TROPONIN C AND
 CC TROPOMYOSIN. THE INTERACTION OF CALPONIN WITH ACTIN INHIBITS THE
 CC ACTOMYOSIN MG-ATPASE ACTIVITY.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Alpha;
 CC IsoId=P26932-1; Sequence=Displayed;


```

DR PRINTS; PR00888; SM22CALPONIN.
DR ProdOm; PD001527; CH type; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS01052; CALPONIN; 3.
DR KAModulin-binding; Actin-binding; Multigene family; Repeat.
FT DOMAIN 28 131
FT REPEAT 164 189 CALPONIN-LIKE 26 AA MOTIF.
FT REPEAT 204 229 CALPONIN-LIKE 26 AA MOTIF.
FT REPEAT 243 268 CALPONIN-LIKE 26 AA MOTIF.
FT CONFLICT 57 57 G -> S (IN REF. 5).
FT CONFLICT 149 149 E -> G (IN REF. 5).
FT CONFLICT 170 170 T -> S (IN REF. 1).
FT CONFLICT 266 266 Q -> P (IN REF. 5).
SQ SEQUENCE 297 AA; 33170 MW; A949573543BC246C CRC64;

Query Match 36.4%; Score 43; DB 1; Length 297;
Best Local Similarity 47.4%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GLAMHTYFPNKGTQOYTQ 19
183 GTRRHLYDPKLGTDQPLDQ 201

RESULT 10
CLP1_MOUSE STANDARD; PRT; 297 AA.
AC 008091;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calponin H1, smooth muscle (Basic calponin) (Calponin 1).
GN CNM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Smooth muscle; PubMed=8370452;
RX MEDLINE=93380560; PubMed=8370452;
RA Strasser P., Glimona M., Moessler H., Herzog M., Small J.V.;
RT "Mammalian calponin. Identification and expression of genetic
variants."
RL FEBS Lett. 330:13-18(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX Gao J.M., Hwang J.M., Resek M.E., Jin J.P.;
RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Liver;
RX Miano J.M., Olson E.N.;
RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX Samaha P.F., Ip H.S., Morrissey E.E., Seltzer J., Tang Z.,
RA Solway J., Parmacek M.S.;
RT "Developmental pattern of expression and genomic organization of the
calponin-h1 gene. A contractile smooth muscle cell marker."
RL J. Biol. Chem. 271:395-403(1996).
RN [5]
RP FUNCTION: THIN FILAMENT-ASSOCIATED PROTEIN THAT IS IMPLICATED IN
THE REGULATION AND MODULATION OF SMOOTH MUSCLE CONTRACTION. IT IS
CAPABLE OF BINDING TO ACTIN, CALMODULIN, TROPONIN C AND
TROPOMYOSIN. THE INTERACTION OF CALPONIN WITH ACTIN INHIBITS THE
ACTOMYOSIN MG-ATPASE ACTIVITY.
CC ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Alpha;
RN [1]

```

```

CC IsoId=Q08091-1; Sequence=Displayed;
CC Name=Beta;
CC IsoId=Q08091-2; Sequence=VSP 00075;
CC TISSUE SPECIFICITY: SMOOTH MUSCLE AND TISSUES CONTAINING
CC SIGNIFICANT AMOUNTS OF SMOOTH MUSCLE.
CC -1 SIMILARITY: BELONGS TO THE CALPONIN FAMILY.
CC -1 SIMILARITY: Contains 1 calponin-homology (CH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z19542; CAA79602.1; -
DR EMBL; I49022; AAB01453.1; -
DR EMBL; I49022; AAB01452.1; -
DR EMBL; U28932; AAC52448.1; -
DR EMBL; U40351; AAC52353.1; -
DR EMBL; U28930; AAC52353.1; JOINED.
DR EMBL; U40348; AAC52353.1; JOINED.
DR EMBL; U40349; AAC52353.1; JOINED.
DR EMBL; U40350; AAC52353.1; JOINED.
DR PIR; S36145; S31486.
DR MGI; MGI:104979; Cnm1.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR000557; Calponin_repeat.
DR InterPro; IPR003247; CH type.
DR InterPro; IPR003096; SM22_calponin.
DR Pfam; PF00402; calponin; 3.
DR Pfam; PR00307; CH; 1.
DR PRINTS; PR00888; SM22CALPONIN.
DR ProDom; PD001527; CH_type; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS01052; CALPONIN; 3.
DR PROSITE; PS50021; CH; 1.
KW Calmodulin-binding; Actin-binding; Multigene family; Repeat;
KW Alternative splicing;
KM DOMAIN 28 131
FT REPEAT 164 189 CH.
FT REPEAT 204 229 CALPONIN-LIKE 26 AA MOTIF.
FT REPEAT 243 268 CALPONIN-LIKE 26 AA MOTIF.
FT VARSPLIC 217 256 Missing (in isoform Beta).
FT CONFLICT 35 35 R -> E (IN REF. 2).
FT CONFLICT 236 237 EH -> GE (IN REF. 4).
SQ SEQUENCE 297 AA; 33355 MW; 0D1F1F31899700ED CRC64;

Query Match 36.4%; Score 43; DB 1; Length 297;
Best Local Similarity 47.4%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GLAMHTYFPNKGTQOYTQ 19
183 GTRRHLYDPKLGTDQPLDQ 201

RESULT 11
CLP1_PIG STANDARD; PRT; 297 AA.
AC 008052;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calponin H1, smooth muscle (Basic calponin) (Calponin 1).
GN CNM1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]

```


GN RFB.
 OS Salmonella borreze.
 OG Plasmid PM0799.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=55400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394842; PubMed=7545154;
 RA Keenleyside W.J., Whitfield C.;
 RT "Lateral transfer of rfb genes: a mobilizable ColE1-type plasmid
 carries the rfB:54 (O:54 antigen biosynthesis) gene cluster from
 RT Salmonella enterica serovar Borreze.";
 RL J. Bacteriol. 177:5247-5253 (1995).
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine = UDP-N-acetyl-D-
 CC mannosamine.
 CC -1- PATHWAY INVOLVED IN THE O:54 LIPOLYSACCHARIDE O-ANTIGEN
 CC BIOSYNTHESIS.
 CC -1- SIMILARITY: TO OTHER UDP-N-ACETYLGLUCOSAMINE 2-EPIMERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L39794; AAC98403.1; -.
 CC HSSP: P27828; 1F6D.
 CC DR InterPro: IPR003331; Epimerase_2.
 CC DR Pfam: PF02350; Epimerase_2; 1.
 CC DR TIGRPFAM: TIGR00236; wecB; 1.
 CC KW Lipopolysaccharide biosynthesis: Isomerase; Plasmid.
 CC SQ SEQUENCE 378 AA; 42495 MW; 0280B9A36CC778C CRC64;
 QY Query Match 36.4%; Score 43; DB 1; Length 378;
 DB Best Local Similarity 38.9%; Pred. No. 22;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 3.NHITPPNKGTOQYTDQI 20
 DB 125 NMNSPPEEGNRQLTSL 142
 RESULT 14
 CH60_PASNU STANDARD; PRT; 547 AA.
 ID CH60_PASNU
 AC 059687;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR GROEL OR MOXA OR PM1107.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 CC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=P1059;
 RX MEDLINE=96105224; PubMed=8529887;
 RA Love B.C., Hansen L.M., Hirsch D.C.;
 RT "Cloning and sequence of the groESL heat-shock operon of Pasteurella
 RT multocida.";
 RL Gene 166:179-180 (1995).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=Pm70;
 CC MEDLINE=21145866; PubMed=11248100;
 CC May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
 CC "Complete genomic sequence of Pasteurella multocida Pm70.";
 CC Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).

CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 CC proper assembly of unfolded polypeptides generated under stress
 CC conditions (By similarity).
 CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 CC 7 subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U30165; AAA84916.1; -.
 CC DR EMBL: AE006151; AAK03191.1; -.
 CC DR PIR: JC4519; JC4519.
 CC DR HSSP: P06139; 1JON.
 CC DR HAMAP: MF 00600; -; 1.
 CC DR InterPro: IPR001844; Chaperin Cpn60.
 CC DR InterPro: IPR002423; Cpn60/TCF-1.
 CC DR Pfam: PF00118; cpn60_TCF1; 1.
 CC DR PRINTS: PR00298; TCAPRONIN60.
 CC DR PRINTS: PR00304; TCOMPLEXCP1.
 CC DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 CC KW Chaperone; ATP-binding; Complete proteome.
 CC FT CONFLICT 424 424 S -> N (IN REF. 1).
 CC FT CONFLICT 464 464 I -> V (IN REF. 1).
 CC FT SEQUENCE 547 AA; 57291 MW; EBB95B9F235B855E CRC64;
 QY Query Match 36.4%; Score 43; DB 1; Length 547;
 DB Best Local Similarity 36.7%; Pred. No. 33;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 10 NKGTOQYTDQIE 21
 DB 479 NAGTEQYGDWIE 490
 RESULT 15
 CH60_HAEIN STANDARD; PRT; 548 AA.
 ID CH60_HAEIN
 AC P43733;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR MOXA OR GROEL OR HI0543.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 CC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shrivley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512 (1995).
 CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
 CC proper assembly of unfolded polypeptides generated under stress
 CC conditions (By similarity).

CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
 CC 7 subunits (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32736; AAC2201.1; -
 CC PIR: C64076; C64076.
 CC HSSP: P06139; 1JON.
 CC TIGR: H10543; -
 CC HAMAP: MF_00600; -; 1.
 CC InterPro: IPR001844; Chaperlin Cpn60.
 CC InterPro: IPR002423; Cpn60/TCP-1.
 CC Pfam: PF00118; Cpn60_TCP1.1.
 CC PRINTS: PR00298; CHAPERONIN60.
 CC PRINTS: PR00304; TCOMPLEXTCP1.
 CC PROSITE: PS00296; CHAPERONINS_CPN60; 1.
 CC Chaperone; ATP-binding; Complete proteome.
 CC KW Chaperone; ATP-binding; Complete proteome.
 CC SQ SEQUENCE 548 AA; 57577 MW; CA406MAC1B62159 CRC64;
 CC -----
 CC Query Match 36.4%; Score 43; DB 1; Length 548;
 CC Best Local Similarity 66.7%; Pred. No. 33;
 CC Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 CC -----
 CC QY 10 NKGTOQYTDQIE 21
 CC Db 480 NAGTEQYGMIE 491
 CC -----
 CC RESULT 16
 CC TR2M AGRVI STANDARD; PRT; 723 AA.
 CC AC 004564;
 CC DT 01-NOV-1995 (Rel. 32, Created)
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
 CC DT 01-NOV-1995 (Rel. 32, Last annotation update)
 CC DE Trypophan 2-monooxygenase (EC 1.13.12.3).
 CC GN IAAW.
 CC OS Agrobacterium vitis (Rhizobium vitis).
 CC OC Plasmid pTIS4.
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 CC NCBI_TaxID=373;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=S4;
 CC RX MEDLINE=93101133; PubMed=1465104;
 CC RA Canada J., Gerard J.-C., Crouzet P., Otten L.;
 CC "Organization and functional analysis of three T-DNAs from the
 CC viopline T1 plasmid pTIS4";
 CC Mol. Gen. Genet. 235:292-303 (1992).
 CC RL -1- CATALYTIC ACTIVITY: L-tryptophan + O(2) = indole-3-acetamide +
 CC CO(2) + H(2)O.
 CC -1- PATHWAY: Biosynthesis of auxins from tryptophan; first step.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M91609; AAA98149.1; -
 CC PIR: S30105; S30105.
 CC InterPro: IPR000759; Adirndx_reductase.

DR InterPro: IPR002937; Amino oxidase.
 DR InterPro: IPR006064; Glycosidase.
 DR InterPro: IPR000205; NAD binding.
 DR InterPro: IPR001878; ZnF_CCHC.
 DR Pfam: PF01593; Amino oxidase; 1.
 DR Pfam: PF02027; Ro1B_Ro1C; 1.
 DR PRINTS: PR00419; ADXRDTASE.
 DR SMART: SM00343; ZnF_C2HC; 1.
 DR Oxidoreductase; Monooxygenase; Auxin biosynthesis; Crown gall tumor;
 DR T-DNA; Plasmid.
 DR KW T-DNA; Plasmid.
 DR SQ SEQUENCE 723 AA; 80803 MW; ABA9CEA94B6B4E36 CRC64;
 CC -----
 CC Query Match 36.4%; Score 43; DB 1; Length 723;
 CC Best Local Similarity 61.5%; Pred. No. 45;
 CC Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 CC -----
 CC QY 1 GINMTYTPNNKGT 13
 CC Db 282 GINMTKEPNNPQT 294
 CC -----
 CC RESULT 17
 CC RS9 MYCTU STANDARD; PRT; 151 AA.
 CC AC 006259;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE 30S ribosomal protein S9.
 CC GN RPS1 OR RV3442C OR MT3547 OR MTCY77.14C.
 CC OS Mycobacterium tuberculosis.
 CC OC Bacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=H37RV;
 CC RX MEDLINE=98295987; PubMed=9634230;
 CC Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 CC Gordon S.V., Eigemeier K., Gae S., Barry C.E. III, Tekaia F.,
 CC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 CC Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 CC Hornsby T., Jagels K., Krogh A., McLean L., Moule S., Murphy L.,
 CC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 CC Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 CC Sultson J.E., Taylor K., Whitehead S., Barrall B.G.;
 CC RT "Deciphering the biology of Mycobacterium tuberculosis from the
 CC RT complete genome sequence."
 CC RL Nature 393:537-544 (1998).
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=CDC 1551 / Oshkosh;
 CC RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 CC Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 CC Kolonay J.F., Nelson W.C., Umayan L., Ermolaeva M.D., Salzberg S.L.,
 CC Bacher A., Uetzelback T., Weidman J., Kouri H., Gill J., Mikula A.,
 CC Bishai W.;
 CC "Whole genome comparison of Mycobacterium tuberculosis clinical and
 CC laboratory strains."
 CC Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z95389; CAB08691.1; -
 CC EMBL: AB007159; AKA4788.1; -
 CC PIR: H70976; H70976.

DR TIGR; M73547; -
 DR Tuberculin; RV3442C; -
 DR HAMAP; MF_00532; -; 1.
 DR InterPro: IPR000754; Ribosomal_S9.
 DR Pfam: PF00380; Ribosomal_S9; 1.
 DR ProDom; PD001627; Ribosomal_S9; 1.
 DR PROSITE; PS00360; RIBOSOMAL_S9; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 151 AA; 16436 MW; 674679A0A1C2FC9E CRC64;

Query Match 36.0%; Score 42.5; DB 1; Length 151;
 Best Local Similarity 35.7%; Pred. No. 9.9;
 Matches 10; Conservative 4; Mismatches 7; Indels 7; Gaps 1;

QY 1 GLNMTYFPNKGTOY-----TDQTE 21
 DB 53 GRSLEDYFPNKHQOLIKAPLVYDRVE 80

RESULT 18
 ID_VIBPA STANDARD; PRT; 228 AA.
 AC 087Q03;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Dethiobiotin synthetase (EC 6.3.3.3) (Dethiobiotin synthase) (DTB
 DE synthetase) (DTBS).
 GN BIOD OR VP1116.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=670;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Nishida M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shitagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 RT distance from that of V. cholerae.";
 RL Lancet 361:743-749(2003).
 CC -1- CATALYTIC ACTIVITY: ATP + 7,8-diaminononanoate + CO(2) = ADP +
 CC phosphate + dethiobiotin.
 CC -1- COPACITOR: Magnesium (By similarity).
 CC -1- PATHWAY: Bioconversion of pimeolate into dethiobiotin.
 CC -1- SIMILARITY: BELONGS TO THE DETHIOBIOTIN SYNTHETASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; APO05076; BACS9379.1; ALT_INT.

DR HAMAP; MF_00335; -; 1.
 KW Biotin biosynthesis; Ligase; Magnesium; ATP-binding;
 KW Complete proteome.
 FT NP BIND 9 17 ATP (BY SIMILARITY).
 SQ SEQUENCE 228 AA; 24711 MW; 9810E29B78E42DF6 CRC64;

Query Match 36.0%; Score 42.5; DB 1; Length 228;
 Best Local Similarity 41.7%; Pred. No. 16;
 Matches 10; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

QY 1 GLNMTYFPNKGTOY-----TDQTE 21
 DB 167 GLNLTGVMIANRVNPGTEHYADITE 190

RESULT 19
 ID_RS9_MYCLE STANDARD; PRT; 153 AA.
 AC P40828;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S9.
 GN RPSI OR M0365 OR B229_C2_191.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrett B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U00020; AAA17296.1; -
 DR EMBL; AL583918; CAC29873.1; -
 DR PIR; S72982; S72982.
 DR Leprosin; M0365; -
 DR HAMAP; MF_00532; -; 1.
 DR InterPro: IPR000754; Ribosomal_S9.
 DR Pfam; PF00380; Ribosomal_S9; 1.
 DR ProDom; PD001627; Ribosomal_S9; 1.
 DR PROSITE; PS00360; RIBOSOMAL_S9; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 153 AA; 16911 MW; B8FD7CD1029F0735 CRC64;

Query Match 35.6%; Score 42; DB 1; Length 153;
 Best Local Similarity 53.3%; Pred. No. 12;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOO 15
 DB 55 GRSLEDYFPNKHQOLIKAPLVYDRVE 69

RESULT 20
 RS9_STRCO STANDARD; PRT; 170 AA.
 ID_RS9_STRCO
 AC Q53875;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S9.
 GN RPSI OR SCO4735 OR SC6G4.13.
 OS Streptomyces coelicolor.

CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Streptomyces; Streptomycetaceae; Streptomyces.
 OK NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=98102746; PubMed=9439573;
 RX Sanchez C., Blanco G., Mendez C., Salas J.A.;
 RT "Cloning, sequencing and transcriptional analysis of a Streptomyces
 RT coellicolor operon containing the rplM and rpsI genes encoding
 RT ribosomal proteins Scoli3 and Scoli5".
 RL Mol. Gen. Genet. 257:91-96(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA MEDLINE=21996410; PubMed=1200953;
 RX Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)".
 RL Nature 417:141-147(2002).
 CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U43429; AAC6061.1; -
 CC EMBL: AL93121; CAA20391.1; -
 CC PIR: T35564; T35564.
 DR HAMAP: MF_00532; -; 1.
 DR InterPro: IPR000754; Ribosomal_S9.
 DR Pfam: PF00380; Ribosomal_S9; 1.
 DR ProDom: PD001627; Ribosomal_S9; 1.
 DR PROSITE: PS00360; RIBOSOMAL_S9; 1.
 RN Ribosomal protein; Complete proteome.
 KM SEQUENCE 170 AA; 1869 MW; 3CEB0FF69151C99 CRC64;
 Query Match 35.6%; Score 42; DB 1; Length 170;
 Best Local Similarity 42.1%; Pred. No. 14;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GINMHTYFPNKGQGYTDQ 19
 DB 71 GRTLEDYFENKVKHOEVNE 89
 RESULT 21
 TRY3_SALSA STANDARD; PRT; 238 AA.
 AC P35033;
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-FEB-2003 (Rel. 42, Last annotation update)
 DE Trypsin III precursor (EC 3.4.21.4) (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Proteanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OK NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;
 RA MEDLINE=96035908; PubMed=7556223;
 RX Male R., Lorens J.B., Smale A.O., Torrisen K.R.;
 RT "Molecular cloning and characterization of anionic and cationic
 RT variants of trypsin from Atlantic salmon."
 RL Eur. J. Biochem. 232:677-685(1995).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X70074; CAA49679.1; -
 CC PIR: S66657; S31779.
 DR PDB: 1AOJ; 13-JAN-99.
 DR MEROPS: S01.151; -
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR SMART: SMO0020; Tryp_Spc; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KM Multigene family; 3D-structure.
 FT FT NON TER 1 7
 FT SIGNAL 1 7
 FT PROPEP 8 15
 FT CHAIN 16 238
 FT ACT_SITE 55 55
 FT ACT_SITE 99 99
 FT ACT_SITE 192 192
 FT DISULFID 22 152
 FT DISULFID 40 56
 FT DISULFID 124 225
 FT DISULFID 131 198
 FT DISULFID 163 177
 FT DISULFID 188 212
 FT SITE 186 186
 SO SEQUENCE 238 AA; 25389 MW; AEF99B80E8399023 CRC64;
 Query Match 35.6%; Score 42; DB 1; Length 238;
 Best Local Similarity 35.0%; Pred. No. 20;
 Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 QY 2 LNMHTYFPNKGQGYTDQ 21
 DB 65 LGEHNIAVNEGTEOPTDSVK 84
 RESULT 22
 CH60_ACTPL STANDARD; PRT; 546 AA.
 AC P94166;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
 GN GROEL OR GROEL OR MOXA.
 OC Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OK NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S 4074 / Serotype 1;
 RA MEDLINE=97189570; PubMed=9037757;
 Vezina G., Strols M., Clatroux N., Boissinot M.;

```

RT "Cloning and characterization of the groE locus from Actinobacillus
RT pleuropneumoniae."
RL FEMS Microbiol. Lett. 147:11-16(1997).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: U55016; AAB51437.1; -.
HASP: P06139; IGR.
HAMAP: MF_00600; -.
DR InterPro: IPR001844; Chaperin Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; Cpn60_TCP1.1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONIN_CPN60. 1.
KW Chaperone; ATP-binding.
FT INIT MET 0
SQ SEQUENCE 546 AA; 57513 MW; E84BB72C9BD3DB56 CRC64;

Query Match 35.6%; Score 42; DB 1; Length 546;
Best Local Similarity 53.3%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 YEPNKGTYQTQDIE 21
DB 476 YGYNAGTEGYDML 490

RESULT 23
Y328_CHLPN STANDARD; PRT; 584 AA.
AC Q9ZBL0; Q9ZJ05;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT Hypothetical protein CPN0328/CP0429/CPJ0328.
CPN0328 OR CP0429 OR CPJ0328.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CML029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lemmel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davila R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
RT Nat. Genet. 21:385-389(1999).
RT [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mofn and Chlamydia
RT pneumoniae AR39";
RT Nucleic Acids Res. 28:1397-1406(2000).

```

```

RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CW1029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SIMILARITY: Belongs to the ubid family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: AE001617; AAD18477.1; -.
DR EMBL: AE002204; AAF38271.1; -.
DR EMBL: AP002546; BAA98538.1; -.
DR PIR: A72092; A72092.
DR PIR: H86531; H86531.
DR TIGR: CP0429; -.
DR InterPro: IPR002830; carboxylase.
DR Pfam: PF01977; Ubid; 1.
DR TIGRPFAM: TIGR00148; TIGR00148. 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 584 AA; 66506 MW; 87F394EE847748F5 CRC64;

Query Match 35.6%; Score 42; DB 1; Length 584;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 HTYPNKGKT 13
DB 573 HAYFNKKT 581

RESULT 24
HS78_CANAL STANDARD; PRT; 812 AA.
AC Q96UX5;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DT Heat shock protein 78, mitochondrial precursor.
GN HSP78.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC5314;
RA Bhatnagar V., Hostetter M.;
RT "The Candida albicans HSP78 gene encoding a member of the ClpB family
RT of stress proteins";
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: SEEMS TO BE INVOLVED IN THE TURNOVER OF UNASSEMBLED
RL MITOCHONDRIAL PROTEINS (By similarity).
CC -1- SUBUNIT: POSSIBLY PART OF AN ATP-DEPENDENT PROTEASE.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; AF399713; AAK97626.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003599; AAA_ATPase_cent.
DR InterPro; IPR001270; Chedrin_c1pA/B.
DR Pfam; PF00004; AAA_1; PROTEINASEA.
DR PRINTS; PR00300; CLPPTOTASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
KW Chaperone; ATP-binding; Repeat; Heat shock; Mitochondrion;
KW Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN 133 140 HEAT SHOCK PROTEIN 78.
FT NP BIND 534 541 ATP (POTENTIAL).
SQ SEQUENCE 812 AA; 91680 MW; 4596001CA6DB7276 CRC64;

Query Match
Best Local Similarity 47.1%; Pred. No. 74;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NMHTYPPNKGTOOYTQD 19
Db 65 NFHSSPRXLQMOQTEQ 81

RESULT 25
RG2 YEAST STANDARD; PRT; 1009 AA.
AC 006407;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Rho-type GTPase-activating protein 2.
GN RGA2 OR YDR379W OR D9481.4.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RC PubMed=9169867;
RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Bahr A., Ballesta J.P.G.,
RA Barques M., Baron L., Becker A., Bileau N., Bloecher H., Blugeon C.,
RA Boskovic J., Brandt P., Bruckner M., Buitrago M.J., Coster F.,
RA Delveau T., del Rey F., Dujon B., Elde L.G., Garcia-Cantalejo J.M.,
RA Goffeau A., Gomez-Peris A., Granotier C., Hamemann V., Hankeln T.,
RA Hohelsel J.D., Jaeger W., Jimenez A., Joniaux J.-L., Kraemer C.,
RA Kuster H., Laamanen P., Legros Y., Louis B.J., Moeller-Rieker S.,
RA Monnet A., Moro M., Mueller-Auer S., Nusshauer B., Paricio N.,
RA Paulin L., Perea J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
RA Prydz H., Purrelle B., Rasmussen S.W., Remacha M., Revuelta J.L.,
RA Rieger M., Salom D., Saluz H.P., Salz J.E., Saren A.-M., Schaefer M.,
RA Schaefer M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
RA Uriestaran L.A., Verhasselt P., Vissers S., Voelt M., Volckaert G.,
RA Wagner G., Wambut R., Wedler E., Wedler H., Wolf S., Harris P.E.,
RA Bowman S., Brown D., Churcher C.W., Connor R., Dedman K., Gentile S.,
RA Hamlin N., Hunt S., Jones L., McDonald S., Murphy L., Niblett D.,
RA Odell C., Oliver K., Rajandream M.A., Richards C., Shore L.,
RA Walsh S.V., Bartell B.G., Dietrich F.S., Mulligan J.T., Allen E.,
RA Araujo R., Aviles E., Berio A., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Hunnicke-Smith S., Hyman R.W., Komp C.,
RA Lahekari D., Lew H., Lin D., Mosedale D., Nakahara K., Natch A.,
RA Oehler P., Oh C., Petel F.X., Roberts D., Schramm S., Schroeder M.,
RA Shogen T., Shroff N., Winant A., Yelton M.A., Botstein D.,
RA Davis R.W., Johnston M., Andrews S., Brinkman R., Cooper J., Ding H.,
RA Du Z., Faville A., Fulton L., Gattung S., Greco T., Hallsworth K.,
RA Hawkins J., Hillier L., Jier M., Johnson D., Johnson L., Kirsten J.,
RA Kucaba T., Langston Y., Latreille P., Le T., Mardis E., Nenezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,
RA Taich A., Treviskis E., Vignati D., Wilcox L., Wohldman P., Vaudin M.,

```

```

RA Wilson R., Waterston R., Albermann K., Hani J., Heumann K., Kleine K.,
RA Mewes H.-W., Zollner A., Zaccaria P.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";
RL Nature 387:75-78(1997).
CC - FUNCTION: GTPASE-ACTIVATING PROTEIN (GAP) FOR CDC42 AND/OR RHOL.
CC - SIMILARITY: Contains 2 LIM zinc-binding domains.
CC - SIMILARITY: Contains 1 Rho-GAP domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U28373; AAB64815.1; -.
DR HSSP; P27986; IPBW.
DR SGD; S0002787; RGA2.
DR InterPro; IPR001781; LIM.
DR InterPro; IPR000198; RhoGAP.
DR Pfam; PF00412; LIM; 2.
DR Pfam; PF00620; RhoGAP; 1.
DR Prodom; PD000094; LIM; 2.
DR SMART; SM00132; LIM; 2.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 1.
DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
DR PROSITE; PS00238; RHO GAP; 1.
KW GTPase activation; Repeat; LIM domain; Metal-binding; Zinc.
FT DOMAIN 11 68 LIM 1.
FT DOMAIN 69 129 LIM 2.
FT DOMAIN 788 1006 RHO-GAP.
SQ SEQUENCE 1009 AA; 113290 MW; FA13BALDF835098 CRC64;

Query Match
Best Local Similarity 38.1%; Pred. No. 93;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 GNMHTYPPNKGTOOYTQD 21
Db 818 GNMGLYKRGSGSQTVEIE 838

RESULT 26
SPB STRPY STANDARD; PRT; 398 AA.
AC P00788; P26296; Q54960; Q54961; Q54962; Q54963; Q54964; Q54965;
AC Q54966; Q54967; Q54968; Q57024; Q57082; Q57202; Q57211; Q57212;
AC Q96880;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Streptolysin precursor (EC 3.4.22.10) (Streptococcal cysteine
DE protease) (Streptococcus peptidase A) (SP) (Exotoxin type B) (SPE
DE B).
GN SPB OR SPY2039 OR SPYM_1742 OR SP51739 OR SPYM18_2039.
OS Streptococcus pyogenes, (serotype M3), and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466, 186103;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-32 AND 146-162.
RC STRAIN=86-858, and NY-5;
RC MEDLINE=90330563; PubMed=2198264;
RX Hauser A.R., Schlievert P.M.;
RT "Nucleotide sequence of the streptococcal pyrogenic exotoxin type B
RT gene and relationship between the toxin and the streptococcal
RT proteinase precursor.";
RN J. Bacteriol. 172:4536-4542(1990).
RN [2]

```

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=789 / Serotype M1, 327 / Serotype M2, 75 / Serotype M3,
RC 157 / Serotype M3, 315 / Serotype M3, 1251 / Serotype M3,
RC 321 / Serotype M4, 1289 / Serotype M5, 303 / Serotype M6,
RC 429 / Serotype M8, 796 / Serotype M9, 800 / Serotype M9,
RC 1896 / Serotype M10, 650 / Serotype M11, 282 / Serotype M12,
RC 659 / Serotype M13, 660 / Serotype M14, 1898 / Serotype M15,
RC 1233 / Serotype M17, 156 / Serotype M18, 300 / Serotype M18,
RC 1294 / Serotype M19, 162 / Serotype M22, 1901 / Serotype M23,
RC 684 / Serotype M24, 686 / Serotype M25, 1838 / Serotype M27,
RC 587 / Serotype M28, 366 / Serotype M30, 427 / Serotype M31,
RC 807 / Serotype M33, 1841 / Serotype M41, 1842 / Serotype M43,
RC 1226 / Serotype M44, 719 / Serotype M49, 1864 / Serotype M56,
RC 1882 / Serotype M59, 168 / Serotype M66, 302 / Serotype M73,
RC 758 / Serotype M75, 1911 / Serotype M75, 1832 / Serotype M76,
RC 165 / Serotype M, 262 / Serotype M, 317 / Serotype M,
RC 1990 / Serotype M, 1991 / Serotype M, 2017 / Serotype M,
RC 2018 / Serotype M, 1719 / Serotype M, 289 / Serotype T28,
RC 1590, 1870, 1871, 1872, 1893, and 1914A;
MEDLINE=94285789; PubMed=7516997;
Kaput V., Topouzis S., Majesky M.W., Li L.L., Hamrick M.R.,
Hamill R.J., Pettit J.M., Musser J.M.;
RT "A conserved Streptococcus pyogenes extracellular cysteine protease
cleaves human fibronectin and degrades vitronectin.";
RT Microb. Pathog. 15:327-346(1993).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=A-20 / Serotype M1, T1;
RC Wu J.-U.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=SV / Serotype M23;
RC Hong K.;
RT "A novel cloning method used arbitrarily primed PCR";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RC MEDLINE=21192684; PubMed=11296286;
RC Ferret J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RC Pirieaux C., Sezane S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RC Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RC Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus
pyogenes.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RC MEDLINE=22133808; PubMed=12122206;
RC Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RC Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RC Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RC Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
phase-encoded toxins, the high-virulence phenotype, and clone
emergence.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
[7]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RC Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RC Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hatotoi M.,
RC Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
of S. pyogenes SSI-1, SF370 and MGAS832.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
[8]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RC MEDLINE=21927593; PubMed=11917108;
RC Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,

RA Sylva G.L., Sturdevant D.B., Rickiela S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kaput V., Dai J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
[9]
RP PRELIMINARY SEQUENCE OF 28-86 AND 121-398.
RA Yonaha K., Elliott S.D., Liu T.-Y.;
RT "Primary structure of streptococcal proteinase.";
RT J. Protein Chem. 1:317-334(1982).
[10]
RP PRELIMINARY SEQUENCE OF 146-398.
RA Tai J.Y., Kott A.A., Liu T.-Y., Elliott S.D.;
RT "Primary structure of streptococcal proteinase. III. Isolation of
cytogen bromide peptides: complete covalent structure of the
polypeptide chain.";
RT J. Biol. Chem. 251:1955-1959(1976).
[11]
RP FUNCTION.
RC STRAIN=VZ131 / Serotype M49, T14;
RC MEDLINE=99081733; PubMed=9864206;
RC Kuo C.-F., Wu J.-U., Tsai P.-J., Kao F.-J., Lei H.-Y., Lin M.T.,
RA Lin Y.-S.;
RT "Streptococcal pyrogenic exotoxin B induces apoptosis and reduces
phagocytic activity in U937 cells.";
RT Infect. Immun. 67:126-130(1999).
[12]
RP FUNCTION.
RC STRAIN=A-20 / Serotype M1, T1;
RC MEDLINE=99386817; PubMed=10456871;
RC Tsai P.-J., Lin Y.-S., Kuo C.-F., Lei H.-Y., Wu J.-U.;
RT "Group A Streptococcus induces apoptosis in human epithelial cells.";
RT Infect. Immun. 67:4334-4339(1999).
[13]
RP FUNCTION: IMPORTANT STREPTOCOCCAL VIRULENCE FACTOR WHICH CLEAVES
HUMAN FIBRONECTIN AND DEGRADES VITRONECTIN. ALSO CLEAVES HUMAN
IL1B PRECURSOR TO FORM BIOLOGICALLY ACTIVE IL1B. CAN INDUCE
APOPTOSIS IN HUMAN MONOCYTES AND EPITHELIAL CELLS IN VITRO, AND
REDUCES PHAGOCYTIC ACTIVITY IN MONOCYTIC CELLS. THUS, MAY PLAY A
ROLE IN BACTERIAL COLONIZATION, INVASION, AND INHIBITION OF WOUND
HEALING.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage with hydrophobic
CC residues at P2, P1 and P1'.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE
CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET
CC FEVER, HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE
CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC
CC FEVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, M66905; AAA26978.1; -
CC EMBL, L26125; AAA26979.1; -
CC EMBL, L26126; AAA26992.1; -
CC EMBL, L26127; AAA26993.1; -
CC EMBL, L26128; AAA26994.1; -
CC EMBL, L26129; AAA26995.1; -
CC EMBL, L26130; AAA26996.1; -
CC EMBL, L26131; AAA26997.1; -
CC EMBL, L26132; AAA26998.1; -
CC EMBL, L26133; AAA26999.1; -
CC EMBL, L26134; AAA27000.1; -
CC EMBL, L26135; AAA27001.1; -
CC EMBL, L26136; AAA27002.1; -

DR EMBL, L26137, AAA27003.1, -
 DR EMBL, L26138, AAA27004.1, -
 DR EMBL, L26139, AAA27005.1, -
 DR EMBL, L26140, AAA27006.1, -
 DR EMBL, L26141, AAA27007.1, -
 DR EMBL, L26142, AAA27008.1, -
 DR EMBL, L26143, AAA27009.1, -
 DR EMBL, L26144, AAA27010.1, -
 DR EMBL, L26145, AAA27011.1, -
 DR EMBL, L26146, AAA27012.1, -
 DR EMBL, L26147, AAA27013.1, -
 DR EMBL, L26148, AAA27014.1, -
 DR EMBL, L26149, AAA27015.1, -
 DR EMBL, L26150, AAA27016.1, -
 DR EMBL, L26151, AAA26980.1, -
 DR EMBL, L26152, AAA26981.1, -
 DR EMBL, L26153, AAA26982.1, -
 DR EMBL, L26154, AAA26983.1, -
 DR EMBL, L26155, AAA26984.1, -
 DR EMBL, L26156, AAA26985.1, -
 DR EMBL, L26157, AAA26986.1, -
 DR EMBL, L26158, AAA26987.1, -
 DR EMBL, L26159, AAA26988.1, -
 DR EMBL, L26160, AAA26989.1, -
 DR EMBL, L26161, AAA26990.1, -
 DR EMBL, L26162, AAA26991.1, -
 DR EMBL, AF104940, AAD17930.1, -

Query Match 35.2%; Score 41.5; DB 1; Length 398;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 LNMHTYFPNKGQOYTD 17
 DB 201 MKTHTY-PNKGLKDYT 215

RESULT 27
 SPEE_BUCAL STANDARD; PRT; 286 AA.
 AC P57305;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Spermidine synthase (EC 2.5.1.16) (Putrescine aminopropyltransferase)
 GN (SPEV).
 OR SPEE OR BU209.
 CC Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Buchnera.
 CC NCBI_TaxID=118099;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN=Tokyo 1998;
 RC MEDLINE=20445173; PubMed=1093077;
 RA Shigenobu S., Watanabe H., Hatori M., Sakai Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- CATALYTIC ACTIVITY: S-adenosylmethioninamine + putrescine = 5'-
 methylthioadenosine + spermidine.
 CC -1- PATHWAY: Spermidine biosynthesis; fifth (last) step.
 CC -1- SIMILARITY: BELONGS TO THE SPERMIDINE/SPERMINE SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL, AP001118, BAB12926.1, -
 DR HAMAP, MF 00198, -; 1.
 DR InterPro, IPR000051, SAM bind.
 DR InterPro, IPR001045, Spermine synthase.
 DR Pfam, PF01564, Spermine synth. 1.
 DR TIGRFAMs, TIGR00417, speB, 1.
 DR PROSITE, PS01330, SPERMIDINE SYNTHASE, 1.
 KW Spermidine biosynthesis; Transferase; Complete proteome.
 FT DOMAIN 79 120
 (POTENTIAL). BINDING TO DECARBOXYLATED SAM
 SQ SEQUENCE 286 AA; 33788 MW; 5EE47D2107EB689E CRC64;

Query Match 34.7%; Score 41; DB 1; Length 286;
 Best Local Similarity 58.3%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YFPNKGTOOYTD 18
 DB 120 YFPNHSNOAYOD 131

RESULT 28
 SM30_MOUSE STANDARD; PRT; 299 AA.
 AC Q64374; Q60944;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
 GN RGN OR SMP30.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6; TISSUE=Liver;
 RC MEDLINE=96328264; PubMed=8765750;
 RA Fujita T., Shirasawa T., Maruyama N.;
 RT "Isolation and characterization of genomic and cDNA clones encoding
 RT mouse senescence marker protein-30 (SMP30).";
 RL Biochim. Biophys. Acta 1308:49-57(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 MEDLINE=97422495; PubMed=9278263;
 RA Murata T., Yamaguchi M.;
 RT "Molecular cloning of the cDNA coding for regucalcin and its mRNA
 RT expression in mouse liver: the expression is stimulated by calcium
 RT administration.";
 RL Mol. Cell. Biochem. 173:127-133(1997).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
 CC IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
 CC CALCIUM SIGNALING IN THE AGED LIVER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: MAINLY PRESENT IN THE LIVER. WEAK EXPRESSION
 CC WAS FOUND IN THE BRAIN AND LUNG, NOT FOUND IN THE KIDNEY.
 CC -1- DEVELOPMENTAL STAGE: PROTEIN AMOUNTS IN LIVER DECREASE
 CC SIGNIFICANTLY WITH AGE.
 CC -1- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U28897, AAC52721.1, -
 DR EMBL, U32170, AAD03478.1, -
 DR EMBL, D86217, BAA13046.1, -
 DR PIR, S72173, S72173.

OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OK NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Fraser A.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolchert G., Aert R., Robben J., Grynoprez B.,
 RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Mosel D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Mambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Kocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usero D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=20309746; PubMed=10748059;
 RA Eitinger T., Degen O., Bohne U., Muller M.;
 RT "Nid1p, a relative of bacterial transition metal permeases in
 RT Schizosaccharomyces pombe, provides nickel ion for urease
 RT biosynthesis";
 RL J. Biol. Chem. 275:18029-18033(2000).
 CC -1- FUNCTION: High-affinity nickel transporter responsible for nickel
 CC uptake. Required for urease biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE HOKN/HUPN/NIXA FAMILY OF NICKEL
 CC TRANSPORTERS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL, AL031787; CA21158.1; -;
 CC EMBL, AL031825; CA21126.1; -;
 CC PIR, T41194; T41194.
 CC GeneDB, Spombe; SPC1884.02; -;
 CC InterPro: IPR004688; NICO.
 CC Pfam: PF03824; NICO.1.
 DR Transport; Nickel; Transmembrane.
 KW TRANSMEM
 FT TRANSMEM 35 57 POTENTIAL.
 FT TRANSMEM 105 127 POTENTIAL.
 FT TRANSMEM 142 164 POTENTIAL.
 FT TRANSMEM 206 228 POTENTIAL.
 FT TRANSMEM 238 255 POTENTIAL.
 FT TRANSMEM 286 308 POTENTIAL.
 FT TRANSMEM 328 347 POTENTIAL.
 FT SEQUENCE 405 AA, 45295 MW, 44CCB1B87D13EFC8 CRC64;

Query Match 34.7%; Score 41; DB 1; Length 405;
 Best Local Similarity 53.8%; Pred. No. 50;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 G1NMETTPNNKKT 13
 DB 346 G1SMHNYFKKKT 358
 RESULT 32
 CD4_SAISC
 ID CD4_SAISC STANDARD; PRT; 457 AA.
 AC Q29037;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
 DE T4/Leu-3).
 GN CD4.
 OS Saimiri sciureus (Common squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
 OK NCBI_TaxID=9521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tatum M., Hashimoto O.;
 RL Submitted (AUG-1996) to the EMBL/Genbank/DDBJ databases.
 CC -1- FUNCTION: ACCESSORY PROTEIN FOR MHC CLASS-II ANTIGEN/T-CELL
 CC RECEPTOR INTERACTION. MAY REGULATE T-CELL ACTIVATION.
 CC -1- SUBUNIT: ASSOCIATES WITH p56-LCK (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 CC EMBL, D86588; BAA1311.1; -;
 CC HSPB; P01730; 1MBR.
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding activity; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR000973; CD4_TcAg.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003596; IG_V.
 DR Pfam: PF00047; IG_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 25
 FT CHAIN 26 457
 FT DOMAIN 26 395
 FT TRANSMEM 396 417
 FT DOMAIN 418 457
 FT DOMAIN 426 457
 FT DOMAIN 126 202
 FT DOMAIN 203 316
 FT DOMAIN 317 373
 FT CARBOHYD 254 254
 FT CARBOHYD 295 295
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 41 109 BY SIMILARITY.
 FT DISULFID 155 184 BY SIMILARITY.
 FT DISULFID 327 369 BY SIMILARITY.
 FT LIPID 418 418 PALMITATE (BY SIMILARITY).
 FT LIPID 421 421 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 457 AA; 50871 MW; 57EBD6344005A015 CRC64;

Query Match 34.7%; Score 41; DB 1; Length 457;
 Best Local Similarity 35.0%; Pred. No. 57;
 Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 2 LNNHTYFPNKGTOQYTDIE 21
 Db 62 LGVQNFVTRGSGKLTDRID 81

RESULT 33
 VLI_HPV19 STANDARD; PRT; 546 AA.

002050;
 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DE 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Major capsid protein L1.
 GN L1.
 OS Human papillomavirus type 19.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus;
 OC NCBI_TaxId=10608;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94265501; PubMed=8205838;
 RA Delius H., Hofmann B.;
 RT "Primer-directed sequencing of human papillomavirus types.";
 RT Curr. Top. Microbiol. Immunol. 186:113-31(1994).
 RN [2]
 RP SEQUENCE OF 344-387 FROM N.A.
 RX MEDLINE=92407963; PubMed=1326639;
 RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;
 RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
 RT variants: a showcase for the molecular evolution of DNA viruses.";
 RL J. Virol. 66:5714-5725(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: X74470; CAA52523.1; -;
 CC EMBL: M96288; AAA47027.1; -;
 CC PIR: S36490; S36490.
 CC InterPro: IPR002210; PV capsid L1.
 CC Pfam: PRO0500; late protein L1.1.
 CC DR PRINTS: PRO0865; HPVcapsidL1.
 CC DR ProDom: PD000544; PV_capsid_L1.1.
 CC KW Coat protein; Late protein.
 CC SQ SEQUENCE 546 AA; 61683 MW; 6E0A86E6640A9CD3 CRC64;

Query Match 34.7%; Score 41; DB 1; Length 546;
 Best Local Similarity 45.0%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

Qy 1 GLNMTYF--PNKGTOQYTD 18
 Db 308 GSKMTYFIPNNSQOQYTN 327

RESULT 34
 HYSA_STRAS

ID HYSA_STRAS STANDARD; PRT; 984 AA.
 AC 053591;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hyaluronate lyase precursor (EC 4.2.2.1) (Hyaluronidase) (Hyase).
 GN HYL OR GBS1270.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus;
 OC NCBI_TaxId=216495;
 OK NCBI_TaxId=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3502 / Serotype III;
 RX MEDLINE=95073999; PubMed=7982914;
 RA Lin B., Hollingshead S.K., Colligan J.E., Egan M.L., Baker J.R.,
 RA Pritchard D.G.;
 RT "Cloning and expression of the gene for group B streptococcal
 RT hyaluronate lyase.";
 RL J. Biol. Chem. 269:30113-30116(1994).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEM316 / Serotype III;
 RX MEDLINE=22242508; PubMed=12354221;
 RA Glaeser P., Ruesch C., Buchrieser C., Chevalier F., Frangoul L.,
 RA Meadek T., Zouine M., Couve E., Lallouet L., Poyart C., Trieu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease.";
 RL Mol. Microbiol. 45:1499-1513(2002).
 CC -1- CATALYTIC ACTIVITY: Hyaluronate = N 3-(4-deoxy-beta-D-gluc-4-
 CC enuronosyl)-N-acetyl-D-glucosamine.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: U15050; AAA56749.1; -;
 CC EMBL: AL766849; CAD46929.1; ALT_INIT.
 CC PIR: A55137; A55137.
 CC PDB: 1PIS; 16-JAN-02.
 CC PDB: 1BQ; 16-JAN-02.
 CC PDB: 1LXM; 30-OCT-02.
 CC DR Sagalier; gbs1270; -;
 CC DR InterPro: IPR003159; Lyase_8.
 CC DR InterPro: IPR004103; Lyase_8_C.
 CC Pfam: PF02278; Lyase_8; 1.
 CC KW Lyase; Signal; Complete proteome; 3D-structure.
 CC KW SIGNAL
 CC FT CHAIN 1
 CC FT SIGNAL 40
 CC FT CHAIN 1
 CC FT CHAIN 41
 CC FT CHAIN 45
 CC FT CHAIN 155
 CC FT CHAIN 183
 CC FT CHAIN 183
 CC FT CHAIN 246
 CC FT CHAIN 267
 CC FT CHAIN 279
 CC FT CHAIN 288
 CC FT CHAIN 292
 CC FT CHAIN 293
 CC FT CHAIN 300
 CC FT CHAIN 300
 CC FT CHAIN 384
 CC FT CHAIN 387
 CC FT CHAIN 413
 CC FT CHAIN 420
 CC FT CHAIN 583
 CC FT CHAIN 609
 CC FT CHAIN 609
 CC FT CHAIN 639
 CC FT CHAIN 666

HAUTURONATE LYASE.
 E -> G (IN REF. 1).
 SA -> FV (IN REF. 1).
 E -> K (IN REF. 1).
 ASTEDK -> GSPEDN (IN REF. 1).
 H -> Y (IN REF. 1).
 IN -> LT (IN REF. 1).
 A -> G (IN REF. 1).
 KT -> EA (IN REF. 1).
 H -> R (IN REF. 1).
 A -> G (IN REF. 1).
 A -> S (IN REF. 1).
 G -> E (IN REF. 1).
 D -> V (IN REF. 1).
 T -> A (IN REF. 1).
 M -> L (IN REF. 1).
 E -> K (IN REF. 1).
 D -> G (IN REF. 1).


```

CC FT CONFLICT 676 676 L -> I (IN REF. 1).
CC FT CONFLICT 688 689 FW -> LG (IN REF. 1).
CC FT CONFLICT 882 882 Q -> L (IN REF. 1).
CC FT CONFLICT 894 894 M -> L (IN REF. 1).
CC SQ SEQUENCE 984 AA; 111580 MW; ACT215BDD5C592 CRC64;

Query Match
Best Local Similarity 34.7%; Score 41; DB 1; Length 984;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 MHTYPPNKGTOOYTDIOE 21
Db 389 MNVFTDAIKYTDPIE 406

RESULT 35
HYSA_STRPN STANDARD; PRT; 1066 AA.
AC 054873; 054874;
15-DEC-1998 (Rel. 37, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
DE Hyaluronate lyase precursor (EC 4.2.2.1) (Hyaluronidase) (Hyase).
GN SP0314.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Terelein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouli H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angluoli S., Dickinson T., Hickey B.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Frazer C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE OF 89-1066 FROM N.A.
RC STRAIN-Type 23;
RX MEDLINE=94156460; PubMed=8112843;
RA Berry A.M., Lock R.A., Thomas S.M., Rajan D.P., Hansman D.,
RA Paton J.C.;
RT "Cloning and nucleotide sequence of the Streptococcus pneumoniae
RT hyaluronidase gene and purification of the enzyme from recombinant
RT Escherichia coli.";
RL Infect. Immun. 62:1101-1108(1994).
RN [3]
RP CRYSSTALLIZATION.
RX MEDLINE=98234706; PubMed=9573623;
RA Jedrejdas M.J., Chantrelat L., Mewbourne R.B.;
RT "Crystallization and preliminary X-ray analysis of Streptococcus
RT pneumoniae hyaluronate lyase.";
CC J. Struct. Biol. 121:73-75(1998).
CC -1- CATALYTIC ACTIVITY: Hyaluronate = N 3-(4-deoxy-beta-D-gluc-4-
CC enuronosyl)-N-acetyl-D-glucosamine.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan
CC by an amide bond (Potential).
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC CC EMBL; AB007344; AAK74491.1; -.
CC DR EMBL; L20670; AAK53685.1; ALT_INIT.
CC DR EMBL; L20670; AAK53686.1; ALT_INIT.
CC DR PIR; B95037; B95037.
CC DR PDB; 1C82; 05-APR-01.
CC DR PDB; 1P9G; 16-MAY-01.
CC DR PDB; 1LOH; 07-AUG-02.
CC DR PDB; 1LXK; 07-AUG-02.
CC DR PDB; 1N7N; 31-DEC-02.
CC DR PDB; 1N7O; 31-DEC-02.
CC DR PDB; 1N7P; 31-DEC-02.
CC DR PDB; 1N7Q; 31-DEC-02.
CC DR PDB; 1N7R; 31-DEC-02.
CC DR TIGR; SP0314; -.
CC DR InterPro; IPR001899; Gram_pos_anchor.
CC DR InterPro; IPR001892; LPXTG.
CC DR InterPro; IPR003159; Lyase_8.
CC DR InterPro; IPR004103; Lyase_8_C.
CC DR Pfam; PF02018; CSM_4_9; 1.
CC DR Pfam; PF00746; Gram_pos_anchor; 1.
CC DR Pfam; PF02278; Lyase_8; 1.
CC DR Pfam; PF02884; Lyase_8_C; 1.
CC DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
CC DR PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE NEG.
CC KW Lyase; Cell wall; peptidoglycan-anchor; signal; Complete proteome;
CC KM 3D-structure.
CC FT SIGNAL 1 30
CC FT CHAIN 31 1039
CC FT PROPEP 1040 1066
CC FT SITE 1036 1040
CC FT MOD_RES 1039 1039
CC FT CONFLICT 108 108
CC FT CONFLICT 115 115
CC FT CONFLICT 139 139
CC FT CONFLICT 211 211
CC FT CONFLICT 236 236
CC FT CONFLICT 290 290
CC FT CONFLICT 313 313
CC FT CONFLICT 340 340
CC FT CONFLICT 613 613
CC FT CONFLICT 658 658
CC FT CONFLICT 821 821
CC FT CONFLICT 848 848
CC FT CONFLICT 853 853
CC FT CONFLICT 907 907
CC SQ SEQUENCE 1066 AA; 120771 MW; 81DB22A837BE61F9 CRC64;

Query Match
Best Local Similarity 34.7%; Score 41; DB 1; Length 1066;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 4 MHTYPPNKGTOOYTDIOE 21
Db 426 MKEYSDDEIKYTDVIE 443

RESULT 36
IF2P_SCHPO STANDARD; PRT; 1079 AA.
ID IF2P_SCHPO
AC Q10251;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable translation initiation factor IF-2.
GN SPAC56F8.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=972;
RX MEDLINE=21698401; PubMed=11659360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodard J., Volkmer G., Aert R., Robben J., Gromprez B.,
RA Welter J., Vancraels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel K., Fuchs M., Filicz C., Holzer E., Moscati D., Halbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mortier S.,
RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
RA Dominguez A., Revelle J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

-1- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING
THE BINDING OF THE FORMYL METHIONINE-TRNA TO RIBOSOMES. SEEMS TO
FUNCTION ALONG WITH EIF-2 (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z69728; CAA93574.1; -
CC PIR; T38913; T38913.
CC DR GeneDB SPombe; SPAC56F8.03; -
CC DR InterPro; IPR000795; EF_Grpbind.
CC DR InterPro; IPR004161; EFTU_D2.
CC DR InterPro; IPR000178; IF2.
CC DR InterPro; IPR005225; Small_GTP.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC PRINTS; PR00315; ELONGATIONFACT.
CC DR TIGRPMG; TIGR00231; small GTP; 1.
CC DR PROSITE; PS01176; IF2; PALISE_NEG.
CC KW Initiation factor; Protein biosynthesis; GTP-binding.
CC FT NP_BIND 491 498 GTP (BY SIMILARITY).
CC SQ SEQUENCE 1079 AA; 119930 MW; 73A01CE933C1F6AA CRC64;

Query Match 34.7%; Score 41; DB 1; Length 1079;
Best Local Similarity 40.0%; Pred. No. 1,4e+02;
Matches 8; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GLNMTYFPNKGQOYTDO 20
Db 655 GLNAAIFYENKNGRIVSLV 674

RESULT 37
DDL LACIA STANDARD; PRT; 349 AA.
AC 09CTLS;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

```

DE D-alanine--D-alanine ligase (EC 6.3.2.4) (D-alanylalanine synthetase)
DE (D-Ala-D-Ala ligase).
GN DDL OR LI0341.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=11403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
RA Weisenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
CC alanyl-D-alanine.
CC -1- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
CC step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE006271; AAK04439.1; -
CC PIR; E86667; E86667.
CC DR HSSP; P07862; IICW.
CC DR HAMAP; MF_00047; -; 1.
CC DR InterPro; IPR005905; D_ala_D_ala.
CC DR InterPro; IPR000291; Dala_lig_Van.
CC Pfam; PF01820; Dala_ligase; 1.
CC DR TIGRPMG; TIGR01205; D ala-D alarIGR; 1.
CC DR PROSITE; PS00843; DALA_DALA_LIGASE_1; 1.
CC DR PROSITE; PS00844; DALA_DALA_LIGASE_2; FALSE_NEG.
CC KW Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
CC SQ SEQUENCE 349 AA; 38693 MW; 8B86996E1D2A8197 CRC64;

Query Match 34.3%; Score 40.5; DB 1; Length 349;
Best Local Similarity 47.4%; Pred. No. 51;
Matches 9; Conservative 3; Mismatches 2; Indels 5; Gaps 1;

Qy 6 TYFPNKG-----TQYTDO 19
Db 40 TYFTKGGEFIKTQFTDK 58

RESULT 38
NIA_CUCMA STANDARD; PRT; 918 AA.
ID NIA_CUCMA
AC P17569;
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrate reductase [NADH] (EC 1.7.1.1) (NR).
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside 1; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1]
RP SEQUENCE FROM N.A.
RA Crawford N.M., Campbell W.H., Davis R.;
RT "Nitrate reductase from squash: cDNA cloning and nitrate regulation.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8073-8076(1986).
RN [2]
RP SEQUENCE FROM N.A.

```

CC TISSUE=Seedling; PubMed=1748631;
 RA MEDLINE=92084635;
 RX Hyde G.B., Crawford N.M., Campbell W.H.;
 RT "The sequence of squash NADH:nitrate reductase and its relationship
 RT to the sequences of other flavoprotein oxidoreductases.";
 RT flavoprotein pyridine nucleotide cytochrome reductases.";
 RL J. Biol. Chem. 266:23542-23547(1991).
 CC -1- FUNCTION: Nitrate reductase is a key enzyme involved in the first
 CC step of nitrate assimilation in plants, fungi and bacteria.
 CC -1- CATALYTIC ACTIVITY: Nitrite + NAD(+) + H(2)O = nitrate + NADH.
 CC -1- COFACTOR: EACH SUBUNIT OF THE ENZYME CONTAINS 1 EQUIVALENT OF FAD,
 CC HEME IRON, AND MOLYBDENUM-PTERIN AS PROSTHETIC GROUPS. THE HEME
 CC GROUP IS CALLED CYTOCHROME B-557.
 CC -1- SUBUNIT: Homodimer.
 CC -1- INDUCTION: By nitrate.
 CC -1- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
 CC N-TERMINAL DOMAIN.
 CC -1- SIMILARITY: Contains 1 cytochrome b5 heme-binding domain.
 CC -1- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
 CC C-TERMINAL DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M33154; AAA33114.1; -
 CC DR PIR, A41667; A41667.
 CC DR HSSP, P17571; 2CND.
 CC DR InterPro: IPR001199; Cyt B5.
 CC DR InterPro: IPR001834; Cyt B5 reductase.
 CC DR InterPro: IPR000572; Euk_Mb_oxred.
 CC DR InterPro: IPR001709; FwM_cyt_reductase.
 CC DR InterPro: IPR005066; Mo-co_dimer.
 CC DR InterPro: IPR001433; Oxred_FAD/NAD(P).
 CC DR Pfam: PF00970; FAD binding_6; 1.
 CC DR Pfam: PF00173; heme_1; 1.
 CC DR Pfam: PF03404; Mo-co_dimer; 1.
 CC DR Pfam: PF00175; NAD_binding_1; 1.
 CC DR Pfam: PF00174; oxidored_molyb; 1.
 CC DR PRINTS: PR00406; CYTB5RDYASE.
 CC DR PRINTS: PR00363; CYTOCHROMEBS.
 CC DR PRINTS: PR00407; EUMOPTERIN.
 CC DR PRINTS: PR00371; FwMCR.
 CC DR ProDom: PD000612; Cyt_B5; 1.
 CC DR ProSITE: PS00191; CYTOCHROME_B5_1; 1.
 CC DR ProSITE: PS00255; CYTOCHROME_B5_2; 1.
 CC DR ProSITE: PS00559; MOLYBDOPTERIN_EUK; 1.
 CC DR Oxidoreductase: Flavoprotein; FAD; NAD; Heme; Molybdenum;
 CC Nitrate assimilation.
 CC KM METAL 195 195 MOLYBDENUM-PTERIN (POTENTIAL).
 CC FT METAL 249 249 MOLYBDENUM-PTERIN (POTENTIAL).
 CC FT METAL 434 434 INTERCHAIN (POTENTIAL).
 CC FT METAL 578 578 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC FT METAL 601 601 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC FT METAL 601 601 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC SQ SEQUENCE 918 AA; 103383 MW; A2CAFDDADEA1B2D1 CRC64;
 CC -----
 CC Query Match 34.3%; Score 40.5; DB 1; Length 918;
 CC Best Local Similarity 47.4%; Pred. No. 1.5e+02;
 CC Matches 9; Conservative 2; Mismatches 7; Indels 1; Gaps 1;
 CC -----
 CC QY 4 MATYPPNKG-TOOYTDDQIE 21
 CC Db 739 VAPKPPNGIMSQYDLSME 757
 CC -----
 CC RESULT 39
 CC RS9_CORGL STANDARD; PRT; 182 AA.
 CC AC Q8NST5;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S9.
 GN RPS1 OR CG10562.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteriales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxId=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RT Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AP005275; BAB97975.1; -
 CC DR HAMAP, MF 00532; -1.
 CC DR InterPro: IPR000754; Ribosomal_S9.
 CC DR Pfam: PF00380; Ribosomal_S9; 1.
 CC DR ProDom: PD001627; Ribosomal_S9; 1.
 CC DR ProSITE: PS00360; RIBOSOMAL_S9; 1.
 CC KM Ribosomal protein; Complete proteome.
 CC SQ SEQUENCE 182 AA; 19665 MW; 69F7D42AC1CE6807 CRC64;
 CC -----
 CC Query Match 33.9%; Score 40; DB 1; Length 182;
 CC Best Local Similarity 53.3%; Pred. No. 30;
 CC Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 CC -----
 CC QY 1 GNMHTYPPNKGTOO 15
 CC Db 84 GRTLEYPNKLHQ 98
 CC -----
 CC RESULT 40
 CC SM30_BOVIN STANDARD; PRT; 299 AA.
 CC ID Q9TUT5;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Senescence marker protein-30 (SMP-30) (Regucalcin) (RC).
 CC GN RGN OR SMP30.
 CC OS Bos taurus (Bovine).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OC Bovidae; Bovinae; Bos.
 CC OX NCBI_TaxId=9913;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Liver;
 CC RX MEDLINE=20351777; PubMed=10891565;
 CC RA Mitsuwa H., Yamaguchi M.;
 CC RT "The gene of Ca2+-binding protein regucalcin is highly conserved in
 CC vertebrate species.";
 CC RL Int. J. Mol. Med. 6:191-196(2000).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF ENZYMATIC ACTIVITY
 CC IN THE LIVER. DECREASE OF SMP30 LEADS TO THE DYSREGULATION OF
 CC CALCIUM SIGNALING IN THE AGED LIVER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SMP-30 / CGR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AB035446; BAA8080.1; -
DR InterPro; IPR005511; SMP-30.
DR Pfam; PF03758; SMP-30; 1.
KW Calcium-binding
SQ SEQUENCE 299 AA; 33308 MW; 4258F7CDB3CCB575 CRC64;

Query Match 33.3%; Score 40; DB 1; Length 299;
Best Local Similarity 26.3%; Pred. No. 52;
Matches 5; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 3 NMHTYPPNKGTOQYTDQIE 21
Db 133 SLVSLFPDHRHVEKYPDQVD 151

Job completed: August 20, 2003, 09:30:37
Job time : 15.1237 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:04 ; Search time 61.0515 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-2
Perfect score: 118
Sequence: 1 GLNMHTYFPNKGTOOYTDQIE 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	138	12	085155 human parvo
2	118	100.0	141	12	085171 human parvo
3	118	100.0	141	12	085166 human parvo
4	118	100.0	141	12	085146 human parvo
5	118	100.0	142	12	085168 human parvo
6	118	100.0	142	12	085173 human parvo
7	118	100.0	144	12	085138 human parvo
8	118	100.0	145	12	085181 human parvo
9	118	100.0	145	12	085161 human parvo
10	118	100.0	146	12	085158 human parvo
11	118	100.0	147	12	085142 human parvo
12	118	100.0	148	12	085150 human parvo
13	118	100.0	151	12	085131 human parvo
14	118	100.0	151	12	085177 human parvo
15	118	100.0	153	12	085123 human parvo
16	118	100.0	157	12	085196 human parvo

17	118	100.0	162	12	085135 human parvo
18	118	100.0	554	12	09PZ59 human parvo
19	118	100.0	554	12	090201 human parvo
20	118	100.0	554	12	065790 human parvo
21	118	100.0	554	12	0912B7 human eryth
22	118	100.0	554	12	09JGP7 human parvo
23	118	100.0	554	12	09WKL9 human parvo
24	118	100.0	554	12	08JN54 human parvo
25	118	100.0	760	12	09PZT8 human parvo
26	118	100.0	761	12	09PZU0 human parvo
27	118	100.0	765	12	09PZT6 human parvo
28	118	100.0	769	12	09PZT4 human parvo
29	118	100.0	781	12	08JN56 human parvo
30	118	100.0	781	12	065789 human parvo
31	118	100.0	781	12	090223 human parvo
32	118	100.0	781	12	085191 human parvo
33	118	100.0	781	12	090221 human parvo
34	118	100.0	781	12	089318 human parvo
35	118	100.0	781	12	090224 human parvo
36	118	100.0	781	12	09PZT0 human parvo
37	118	100.0	781	12	089320 human parvo
38	118	100.0	781	12	085117 human parvo
39	118	100.0	781	12	089321 human parvo
40	118	100.0	781	12	090250 human parvo
41	118	100.0	781	12	090222 human parvo
42	118	100.0	781	12	0912B8 human eryth
43	118	100.0	781	12	09WKM0 human parvo
44	118	100.0	781	12	089319 human parvo
45	118	100.0	781	12	090200 human parvo

ALIGNMENTS

RESULT 1

ID 085155 PRELIMINARY; PRT; 138 AA.

AC 085155; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VPI.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;

RN [1]

RP SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotski A., Giegler A., Cassinotti P., Siegl G., Wolf H., Modrow S.;

RT "XXXXSequence variability among different parvovirus B19 isolates."

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70564; CAA94477.1; -.

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1

FT NON_TER 138

SQ SEQUENCE 138 AA; 15237 MW; 5899PB8879A3B6BD CRC64;

Query Match 100.0%; Score 118; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
Db 98 GLNMHTYFPNKGTOOYTDQIE 118

RESULT 2

ID 085171 PRELIMINARY; PRT; 141 AA.

AC 085171; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70580; CA94493.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT NON_TER 1 141
 SQ SEQUENCE 141 AA; 15770 MW; C9B92572A78C6C27 CRC64;
 QY Query Match 100.0%; Score 118; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GLNMTYFPNKGTOQYTDQIE 21
 100 GLNMTYFPNKGTOQYTDQIE 120

RESULT 3
 ID 085166 PRELIMINARY; PRT; 141 AA.
 AC 085166;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70575; CA94488.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT NON_TER 1 141
 SQ SEQUENCE 141 AA; 15785 MW; 4115D3D915751757 CRC64;

QY Query Match 100.0%; Score 118; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GLNMTYFPNKGTOQYTDQIE 21
 100 GLNMTYFPNKGTOQYTDQIE 120

RESULT 4
 ID 085146 PRELIMINARY; PRT; 141 AA.
 AC 085146;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70555; CA94467.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT NON_TER 1 141
 SQ SEQUENCE 141 AA; 15712 MW; 32FEB23E7E6B8583 CRC64;
 QY Query Match 100.0%; Score 118; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GLNMTYFPNKGTOQYTDQIE 21
 97 GLNMTYFPNKGTOQYTDQIE 117

RESULT 5
 ID 085168 PRELIMINARY; PRT; 142 AA.
 AC 085168;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,
 Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70577; CA94490.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 142
 FT NON_TER 1 142
 SQ SEQUENCE 142 AA; 15821 MW; 2B6E4D9A6784F8C4 CRC64;
 QY Query Match 100.0%; Score 118; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GLNMTYFPNKGTOQYTDQIE 21
 96 GLNMTYFPNKGTOQYTDQIE 116

RESULT 6
 ID 085173 PRELIMINARY; PRT; 142 AA.
 AC 085173;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,

RA Wolf H., Modrow S.;
 RT "XXSsequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70582; CAA94495.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 142
 FT NON_TER 1 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A95E07C0BC0434 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 118; DB 12; Length 142;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
 DB 96 GINMHTYFPNKGTOOYTDOIE 116

RESULT 7
 085138 PRELIMINARY; PRT; 144 AA.

AC 085138;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemner A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSsequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70547; CAA94459.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 144
 FT NON_TER 1 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6E37BA5997AE CRC64;

Query Match
 Best Local Similarity 100.0%; Score 118; DB 12; Length 144;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GINMHTYFPNKGTOOYTDOIE 21
 98 GINMHTYFPNKGTOOYTDOIE 118

RESULT 8
 085181 PRELIMINARY; PRT; 145 AA.

AC 085181;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemner A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSsequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70550; CAA94503.1; -
 FT NON_TER 1 146
 FT NON_TER 1 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FB95AB6E4D9A CRC64;

FT NON_TER 1 145
 FT NON_TER 1 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1E1B9B923C4 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 118; DB 12; Length 145;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
 DB 100 GINMHTYFPNKGTOOYTDOIE 120

RESULT 9

085161 PRELIMINARY; PRT; 145 AA.

AC 085161;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemner A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSsequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70570; CAA94483.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 145
 FT NON_TER 1 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 118; DB 12; Length 145;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
 DB 96 GINMHTYFPNKGTOOYTDOIE 116

RESULT 10
 085158 PRELIMINARY; PRT; 146 AA.

AC 085158;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemner A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXSsequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70567; CAA94480.1; -
 DR InterPro: IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 146
 FT NON_TER 1 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FB95AB6E4D9A CRC64;

Query Match
 Best Local Similarity 100.0%; Score 118; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
DB 96 GLNMHTYFPNKGTOOYTDQIE 116

RESULT 11

Q85142 PRELIMINARY; PRT; 147 AA.

AC Q85142; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270551; CAA94463.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 147
FT NON_TER 147 147

QY SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
DB 99 GLNMHTYFPNKGTOOYTDQIE 119

RESULT 12

Q85150 PRELIMINARY; PRT; 148 AA.

AC Q85150; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270559; CAA94471.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 148
FT NON_TER 148 148

QY SEQUENCE 148 AA; 16539 MW; 8814ECF2459B308B CRC64;

Query Match 100.0%; Score 118; DB 12; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
DB 100 GLNMHTYFPNKGTOOYTDQIE 120

DB 99 GLNMHTYFPNKGTOOYTDQIE 119

RESULT 13

Q85131 PRELIMINARY; PRT; 151 AA.

AC Q85131; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "Sequence variability among different parvovirus B19 isolates."
DL EMBL; 270540; CAA94452.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 151
FT NON_TER 151 151

QY SEQUENCE 151 AA; 16902 MW; 3CA74914B8E72A3E CRC64;

Query Match 100.0%; Score 118; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
DB 100 GLNMHTYFPNKGTOOYTDQIE 120

RESULT 14

Q85177 PRELIMINARY; PRT; 151 AA.

AC Q85177; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
RT "XXXSequence variability among different parvovirus B19 isolates."
DL EMBL; 270586; CAA94499.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 151
FT NON_TER 151 151

QY SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOOYTDQIE 21
DB 100 GLNMHTYFPNKGTOOYTDQIE 120

RESULT 15


```

085123. PRELIMINARY; PRT; 153 AA.
ID 085123;
AC 085123;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DE 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332516; PubMed=8760426;
RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
RT "Sequence variability among different parvovirus B19 isolates.";
RL J. Gen. Virol. 77:1781-1785(1996).
DR EMBL; 270532; CA94444.1; -.
InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 153
SQ SEQUENCE 153 AA; 17131 MW; ECCALF44020814EC CRC64;

Query Match
Best Local Similarity 100.0%; Score 118; DB 12; Length 153;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNMHTYFPNKGTOOYTDOIE 21
Db 99 GNNMHTYFPNKGTOOYTDOIE 119

RESULT 16
ID 085196 PRELIMINARY; PRT; 157 AA.
AC 085196;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; 270603; CA94518.1; -.
InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 157
SQ SEQUENCE 157 AA; 17591 MW; 1D42191887FFCE03 CRC64;

Query Match
Best Local Similarity 100.0%; Score 118; DB 12; Length 157;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNMHTYFPNKGTOOYTDOIE 21
Db 101 GNNMHTYFPNKGTOOYTDOIE 121

RESULT 17
ID 085135 PRELIMINARY; PRT; 162 AA.
AC 085135;
DT 01-NOV-1996 (TREMblrel. 01, Created)

```

```

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
  Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; 270544; CA94456.1; -.
InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 162
SQ SEQUENCE 162 AA; 18005 MW; 3F51443566660F2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 118; DB 12; Length 162;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNMHTYFPNKGTOOYTDOIE 21
Db 100 GNNMHTYFPNKGTOOYTDOIE 120

RESULT 18
ID 09PZS9 PRELIMINARY; PRT; 554 AA.
AC 09PZS9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VP2 capsid protein.
GN VP.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HV;
RA Gallinella G., Ventunoli S.;
RT "B19 Genome Sequence and Structure Analysis.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF162273; AAD46615.1; -.
InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60833 MW; 7901FBD65A697B6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 118; DB 12; Length 554;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNNMHTYFPNKGTOOYTDOIE 21
Db 386 GNNMHTYFPNKGTOOYTDOIE 406

RESULT 19
ID 090201 PRELIMINARY; PRT; 554 AA.
AC 090201;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE VP2 structural protein (Capsid protein VP2) (Fragment).
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97081188; PubMed=8922470;
RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
  VP1/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M1, and N8;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K.,
  Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with
  rheumatoid arthritis.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=N8;
RX MEDLINE=90218047; PubMed=2157807;
  Umeno K., Nunoue T.;
  "The genome type of human parvovirus B19 strains isolated in Japan
  during 1981 differs from types detected in 1986 to 1987: a correlation
  between genome type and prevalence.";
RL J. Gen. Virol. 71:983-986(1990).
DR EMBL; U53595; AAB47453.1; -
DR EMBL; U53596; AAB47455.1; -
DR EMBL; U53593; AAB47449.1; -
DR EMBL; U53594; AAB47451.1; -
DR EMBL; U53597; AAB47457.1; -
DR EMBL; U53600; AAB47463.1; -
DR EMBL; U53601; AAB47465.1; -
DR EMBL; U53598; AAB47459.1; -
DR EMBL; U53599; AAB47461.1; -
DR EMBL; AB030693; BAA90290.1; -
DR EMBL; AB030673; BAA90268.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 554
SQ SEQUENCE 554 AA; 60853 MW; 98F598FF20CB66F CRC64;

Query Match 100.0%; Score 118; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUNMHTYFPNKGTOOYTDOIE 21
DB 386 GUNMHTYFPNKGTOOYTDOIE 406

RESULT 20
ID 065790 PRELIMINARY; PRT; 554 AA.
AC 065790;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
RT VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VEN1;
RA Echevarria Mayo J.E., Erdman D.D.;
RT Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U31358; AAA83559.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60798 MW; BA89F2B293BE4E24 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GUNMHTYFPNKGTOOYTDOIE 21
DB 386 GUNMHTYFPNKGTOOYTDOIE 406

RESULT 21
ID 0912B7 PRELIMINARY; PRT; 554 AA.
AC 0912B7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE VP2 protein.
GN VP2.
OS Human erythrovirus V9.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=72197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V9;
RA Nguyen Q.T.;
RT "Molecular cloning and sequencing of a novel human erythrovirus
  genome: new species beside B19 in the genus Erythrovirus.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249437; CAC80622.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60885 MW; BA738CDDA80F36 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUNMHTYFPNKGTOOYTDOIE 21
DB 386 GUNMHTYFPNKGTOOYTDOIE 406

RESULT 22
ID 09JGP7 PRELIMINARY; PRT; 554 AA.
AC 09JGP7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Capsid protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rm;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K.,
  Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with
  rheumatoid arthritis.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030694; BAA90293.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60839 MW; 0DB950B33C73EF64 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUNMHTYFPNKGTOOYTDOIE 21
DB 386 GUNMHTYFPNKGTOOYTDOIE 406

```

RESULT 23
Q9WKL9 PRELIMINARY; PRT; 554 AA.
AC Q9WKL9; 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Structural protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20014169; PubMed=10548133;
RA Hemeny A., Beckenlehner K., Wolf H., Lang B., Modrow S.;
RT "Acute parvovirus B19 infection in connection with a flare of systemic lupus erythematosus in a female patient.";
J. Clin. Virol. 14:73-77(1999).
EMBL; AF113323; AAC99439.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60913 MW; 98FB588F205C66F CRC64;
Query Match 100.0%; Score 118; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GINMHTYFPNKGTOQYTDQIE 21
Db 386 GINMHTYFPNKGTOQYTDQIE 406
RESULT 24
Q8JNS4 PRELIMINARY; PRT; 554 AA.
AC Q8JNS4; 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=D91.1;
MEDLINE=22174902; PubMed=12186896;
RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G.,
RT "Genetic diversity within Human Erythroviruses: Identification of three genotypes.";
J. Virol. 76:9124-9134(2002).
EMBL; AY083234; AAL91014.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;
Query Match 100.0%; Score 118; DB 12; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GINMHTYFPNKGTOQYTDQIE 21
Db 386 GINMHTYFPNKGTOQYTDQIE 406
RESULT 25
Q9PZT8 PRELIMINARY; PRT; 760 AA.
AC Q9PZT8; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=Kat12;
MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in human synovial tissue.";
J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161224; AAD45912.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
FT NON_TER 760
SQ SEQUENCE 760 AA; 83403 MW; EA6EB0145E3A0E5A CRC64;
Query Match 100.0%; Score 118; DB 12; Length 760;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GINMHTYFPNKGTOQYTDQIE 21
Db 613 GINMHTYFPNKGTOQYTDQIE 633
RESULT 26
Q9PZU0 PRELIMINARY; PRT; 761 AA.
AC Q9PZU0; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=Kat11;
MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in human synovial tissue.";
J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161223; AAD45910.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
FT NON_TER 761
SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;
Query Match 100.0%; Score 118; DB 12; Length 761;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GINMHTYFPNKGTOQYTDQIE 21
Db 613 GINMHTYFPNKGTOQYTDQIE 633
RESULT 27
Q9PZT6 PRELIMINARY; PRT; 765 AA.
AC Q9PZT6; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VP1/2 (Fragment).
OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati3;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokyar K., Brunstein J., Soderlund-Venemo M., Kivluoto O.,
RT Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL: AF161225; AAD45915.1; -;
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON TER 765 765
SQ SEQUENCE 765 AA; 83998 MW; 89E2546086DCDB8 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 765;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GLNMTYFPNKGTOQYTDQIE 21
613 GLNMTYFPNKGTOQYTDQIE 633

RESULT 28
Q9PZT4 PRELIMINARY; PRT; 769 AA.
AC Q9PZT4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kati4;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokyar K., Brunstein J., Soderlund-Venemo M., Kivluoto O.,
RT Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL: AF161226; AAD45917.1; -;
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
FT NON TER 769 769
SQ SEQUENCE 769 AA; 84578 MW; 0749D46E5CA7BB68 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 769;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQIE 21
613 GLNMTYFPNKGTOQYTDQIE 633

RESULT 29
Q8JNS6 PRELIMINARY; PRT; 781 AA.
AC Q8JNS6;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=D91.1;
RX MEDLINE=22174902; PubMed=12186896;
RA Servant A., Lapereche S., Lallemand F., Marinho V., De Saint Maur G.,
RT Werlic J.F., Garday-Giron A.;
RT "Genetic Diversity within Human Erythroviruses: Identification of
Three Genotypes.";
RL J. Virol. 76:9124-9134(2002).
DR EMBL: AY083234; AAL91013.1; -;
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86142 MW; 4EB71FEDD41FC8F3 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQIE 21
613 GLNMTYFPNKGTOQYTDQIE 633

RESULT 30
Q65789 PRELIMINARY; PRT; 781 AA.
AC Q65789;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VEN1;
RA Echevarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U31358; AAA83558.1; -;
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMTYFPNKGTOQYTDQIE 21
613 GLNMTYFPNKGTOQYTDQIE 633

RESULT 31
P90223 PRELIMINARY; PRT; 781 AA.
AC P90223;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VPI and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA2;
RX MEDLINE=97081188; PubMed=8922470;
RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
VPI/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL: U38514; AAB47796.1; -;
DR InterPro: IPR001403; Parvo_coat.

DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; CID911E39CF04A88 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
DB 613 GINMHTYFPNKGTOOYTDOIE 633

RESULT 32

ID 085191 PRELIMINARY; PRT; 781 AA.

AC 085191;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
ORF2 protein.

OC Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OK NCBI_TaxID=10798;

RN [1]
RA SEQUENCE FROM N.A.
RA Hemeny A., Von Pohlitzki A., Giegler A., Cassinotti P., Siegl G.,
RA Wolf H., Modrow S.;

RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (Apr-1996) to the EMBL/Genbank/DBD databases.

DR EMBL; 270599; CAA94513.1; -
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86026 MW; AA02577B683EDB2D CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
DB 613 GINMHTYFPNKGTOOYTDOIE 633

RESULT 33

ID P90221 PRELIMINARY; PRT; 781 AA.

AC P90221;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.

OC Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OK NCBI_TaxID=10798;

RN [1]
RA SEQUENCE FROM N.A.

RC STRAIN=KOR2;
RX MEDLINE=97081188; PubMed=8922470;
RX Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774 (1996).

DR EMBL; U38511; AAB47793.1; -
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; 9FA830083F6F1357 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21

DB 613 GINMHTYFPNKGTOOYTDOIE 633

RESULT 34

ID P89318 PRELIMINARY; PRT; 781 AA.

AC P89318;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.

OC Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OK NCBI_TaxID=10798;

RN [1]
RA SEQUENCE FROM N.A.

RC STRAIN=JAP1;
RX MEDLINE=97081188; PubMed=8922470;
RX Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774 (1996).

DR EMBL; U38509; AAB47791.1; -
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86064 MW; 3AE65CE69096339 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
DB 613 GINMHTYFPNKGTOOYTDOIE 633

RESULT 35

ID P90224 PRELIMINARY; PRT; 781 AA.

AC P90224;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.

OC Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OK NCBI_TaxID=10798;

RN [1]
RA SEQUENCE FROM N.A.

RC STRAIN=BRZ1;
RX MEDLINE=97081188; PubMed=8922470;
RX Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774 (1996).

DR EMBL; U38546; AAB47801.1; -
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86039 MW; C3A29EB1DB8AD378 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GINMHTYFPNKGTOOYTDOIE 21
DB 613 GINMHTYFPNKGTOOYTDOIE 633

RESULT 36

ID Q9PZT0 PRELIMINARY; PRT; 781 AA.

AC Q9PZT0; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP1 capsid protein.
GN VP.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HV;
RA Gallinella G., Venturoli S.;
RT "B19 Genome Sequence and Structure Analysis."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; ARL62273; AAD46514.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86020 MW; 225DF1EDC68B841F CRC64;
Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOQYTDQIE 21
DB 613 GLNMHTYFPNKGTOQYTDQIE 633

RESULT 37
P89320 PRELIMINARY; PRT; 781 AA.
ID P89320;
AC P89320;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA4;
RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the VP1/VP2 gene from multiple isolates."
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; U38516; AAB47798.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86053 MW; C3A256D548A7DA4 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOQYTDQIE 21
DB 613 GLNMHTYFPNKGTOQYTDQIE 633

RESULT 38
P89317 PRELIMINARY; PRT; 781 AA.
ID P89317;
AC P89317; O12798; O12505;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxId=10798;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=STU;
RX MEDLINE=96370794; PubMed=8774690;
RA Hicks K.E., Cubel R.C., Cohen B.J., Clewley J.P.;
RT "Sequence analysis of a parvovirus B19 isolate and baculovirus expression of the non-structural protein."
RL Arch. Virol. 141:1319-1327(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=USA1;
RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the VP1/VP2 gene from multiple isolates."
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; Z68146; CAA32270.1; -
DR EMBL; U38513; AAB47795.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86041 MW; 4437DAF558F6C752 CRC64;
Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOQYTDQIE 21
DB 613 GLNMHTYFPNKGTOQYTDQIE 633

RESULT 39
P89321 PRELIMINARY; PRT; 781 AA.
ID P89321;
AC P89321;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA5;
RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the VP1/VP2 gene from multiple isolates."
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; U38517; AAB47799.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86056 MW; 92B618F93D2A85D1 CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLNMHTYFPNKGTOQYTDQIE 21
DB 613 GLNMHTYFPNKGTOQYTDQIE 633

RESULT 40
Q9JGSO PRELIMINARY; PRT; 781 AA.
ID Q9JGSO;
AC Q9JGSO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein VP1.

OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K.,
 RA Sasaki T.;
 RT "Sequence of human parvovirus B19 isolates from patients with
 RT rheumatoid arthritis."
 RU Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RX MEDLINE=90218047; PubMed=2157807;
 RA Umene K., Nunoue T.;
 RT "The genome type of human parvovirus B19 strains isolated in Japan
 RT during 1981 differs from types detected in 1986 to 1987: a correlation
 RT between genome type and prevalence."
 J. Gen. Virol. 71:983-986(1990).
 EMBL; AB030673; BAA90267.1; -.
 InterPro: IPR001403; Parvo_coat.
 Pfam: PF00740; Parvo_coat_1.
 DR Pfam: PF00740; Parvo_coat_1.
 SQ SEQUENCE 781 AA; 86026 MW; C9PDP3254421693A CRC64;

Query Match 100.0%; Score 118; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GINMHTYPPNKGTOOYTDQIE 21
 ||||||||||||||||
 Db 613 GINMHTYPPNKGTOOYTDQIE 633

Search completed: August 20, 2003, 09:23:57
 Job time : 63.0515 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:04 ; Search time 40.8247 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-4

Perfect score: 64
Sequence: 1 NKGTYQYTDQIE 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq 19Jun03:*

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	370	12	AA13406
2	64	100.0	543	12	AA13405
3	64	100.0	554	16	AAW08987
4	64	100.0	554	20	AAV23230
5	64	100.0	554	24	ABP57264
6	64	100.0	554	24	ABP57267
7	64	100.0	756	21	AAV71231
8	64	100.0	781	16	AAW08986
9	64	100.0	781	20	AAV23227

10	64	100.0	781	24	ABP57263	Human parvovirus B
11	64	100.0	781	24	ABP57266	Human parvovirus B
12	41	64.1	237	22	ABB60093	Drosophila melanog
13	40	62.5	187	20	AAV14929	Amino acid sequenc
14	39	60.9	277	22	ABB64125	Drosophila melanog
15	38	59.4	522	23	AAO15899	Pseudomonas glutar
16	38	59.4	522	23	AAO15900	Mutant Pseudomonas
17	38	59.4	522	23	AAO15901	Mutant Pseudomonas
18	38	59.4	522	23	AAO15902	Mutant Pseudomonas
19	38	59.4	522	23	AAO15903	Mutant Pseudomonas
20	38	59.4	522	23	AAO15904	Mutant Pseudomonas
21	38	59.4	522	23	AAO15905	Mutant Pseudomonas
22	38	59.4	522	23	AAO15906	Mutant Pseudomonas
23	38	59.4	522	23	AAO15907	Mutant Pseudomonas
24	38	59.4	522	23	AAO15908	Mutant Pseudomonas
25	38	59.4	522	23	AAO15910	Mutant Pseudomonas
26	38	59.4	522	23	AAO15913	Mutant Pseudomonas
27	38	59.4	522	23	AAO15916	Mutant Pseudomonas
28	38	59.4	522	23	AAO15915	Mutant Pseudomonas
29	38	59.4	522	23	AAO15916	Mutant Pseudomonas
30	38	59.4	522	23	AAO15917	Mutant Pseudomonas
31	38	59.4	522	23	AAO15918	Mutant Pseudomonas
32	38	59.4	522	23	AAO15919	Mutant Pseudomonas
33	38	59.4	522	23	AAO15920	Mutant Pseudomonas
34	38	59.4	522	23	AAO15921	Mutant Pseudomonas
35	38	59.4	522	23	AAO15922	Mutant Pseudomonas
36	38	59.4	522	23	AAO15923	Mutant Pseudomonas
37	38	59.4	522	23	AAO15924	Mutant Pseudomonas
38	38	59.4	522	23	AAO15925	Mutant Pseudomonas
39	38	59.4	522	23	AAO15926	Mutant Pseudomonas
40	38	59.4	522	23	AAO15927	Mutant Pseudomonas
41	38	59.4	522	23	AAO15928	Mutant Pseudomonas
42	38	59.4	522	23	AAO15929	Mutant Pseudomonas
43	38	59.4	522	23	AAO15930	Mutant Pseudomonas
44	38	59.4	522	23	AAO15931	Mutant Pseudomonas
45	38	59.4	522	23	AAO15932	Mutant Pseudomonas

ALIGNMENTS

RESULT 1
AA13406
AA13406 standard; Protein; 370 AA.

AC AA13406;
24-OCT-1991 (first entry)
XX Parvo virus B19 PANSE.
DE
XX Primer; PCR; PAPST; globulin.
XX
OS Synthetic.
XX
PN DE4003826-A.
XX
PD 14-AUG-1991.
XX
XX 08-FEB-1990; 90DE-4003826.
XX
XX 08-FEB-1990; 90DE-4003826.
XX
XX 08-FEB-1990; 90DE-4003826.
XX
XX 08-FEB-1990; 90DE-4003826.
XX
XX (MIKR-) MIKROGEN MOLEKULARB.
XX
XX Soutschek E, Motz M;
XX WPI; 1991-246423/34.
XX
XX Immunologically active parvo virus B19 peptide(s) - comprising
XX capsid protein VP1 or VP2 fragments, useful for antibody
XX detection or vaccination
XX

PS Claim 16; Page 10-11; 22pp; German.
 CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAQ13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 XX
 SQ Sequence 370 AA;

Query Match 100.0%; Score 64; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. No. 0.0019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQIE 12
 |||||
 231 NKGTQOYTDQIE 242

RESULT 2
 AAR13405
 ID AAR13405 standard; Protein; 543 AA.
 XX
 AC AAR13405;

DT 24-OCT-1991 (first entry)

DE Parvo virus B19 VP2.

KX Primer; PCR; globulin; PANSE; PAPST.

XX Synthetic.

PN DE4003826-A.

XX 14-AUG-1991.

PF 08-FEB-1990; 90DE-4003826.

PR 08-FEB-1990; 90DE-4003826.

PA (MIKR-) MIKROGEN MOLEKULARB.

XX Soutschek E, Motz M;

XX WPI; 1991-246423/34.

PT Immunologically active parvo virus B19 peptide(s) - comprising
 PT capsid protein VP1 or VP2 fragments, useful for antibody
 PT detection or vaccination

PS Disclosure; Fig 2-6; 22pp; German.

XX VP2 and its fragments PANSE (AAR13406) and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAQ13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 XX

SQ Sequence 543 AA;

Query Match 100.0%; Score 64; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQIE 12
 |||||
 Db 404 NKGTQOYTDQIE 415

RESULT 3
 AAM08987
 ID AAM08987 standard; Protein; 554 AA.
 XX
 AC AAM08987;

DT 27-FEB-1997 (first entry)

DE Human parvovirus VP-2 protein.

KW Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KW non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KW erythblastemia; abortion; universal fetal hydrops; liver disease;
 KW haemorrhagic fever; rheumatism; detection; IgG antibody.

OS Human parvovirus.

PN JP07147986-A.

PD 13-JUN-1995.

PF 24-SEP-1992; 92JP-0281017.

PR 24-SEP-1992; 92JP-0281017.

PA (DENK-) DENKA SEIKEN KK.

PA (ELED-) DENKI KAGAKU KOGYO KK.

DR WPI: 1995-242756/32.

DR N-PSDB; AAT49535.

PT Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.

PS Claim 3; Page 7-9; 38pp; English.

XX The sequences given in AAM08986 represent the parvovirus structural
 CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.
 XX

SQ Sequence 554 AA;

Query Match 100.0%; Score 64; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQIE 12
 |||||
 Db 395 NKGTQOYTDQIE 406

RESULT 4
 AAY23230
 ID AAY23230 standard; Protein; 554 AA.

XX AAY23230;

DT 26-AUG-1999 (first entry)

DE Erythrovirus V9 VP2 protein.

KW Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KW erythrovirus screening; typing; immunoassay; VP2 protein.

XX Erythrovirus.
 OS FR2771751-A1.
 XX
 XX 04-JUN-1999.
 XX
 XX 03-DEC-1997; 97FR-0015197.
 XX
 XX 03-DEC-1997; 97FR-0015197.
 XX
 XX (ASSI-) ASSISTANCE PUBLIQUE HOPITALUX PARIS.
 XX
 XX Auguste V, Garbary CA, Nguyen QT;
 XX WPI; 1999-349543/30.
 XX N-PSDB; ABX81586.
 XX
 XX Erythrovirus V9 and its nucleic acid sequences - can be used in the
 PT diagnosis of its infections
 XX
 XX Claim 19; Page 57-58; 80pp; French.
 XX
 XX The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AA81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.
 XX

Sequence 554 AA;
 SQ
 Query Match 100.0%; Score 64; DB 20; Length 554;
 Best Local Similarity 100.0%; Pred. NO. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NKGTOQYTDQIE 12
 |||||
 Db 395 NKGTOQYTDQIE 406

RESULT 5
 ABP57264
 ID ABP57264 standard; Protein; 554 AA.
 XX
 XX ABP57264;
 XX
 XX 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.
 XX
 KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 XX
 PR 19-MAR-2002; 2002US-365956P.
 XX
 PR 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX
 DR WPI; 2003-201510/19.
 XX
 DR N-PSDB; ABX59574.
 PT

XX Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -
 XX
 XX Example 4; Fig 7B; 148pp; English.
 XX

The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to
 CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX

Sequence 554 AA;
 SQ
 Query Match 100.0%; Score 64; DB 24; Length 554;
 Best Local Similarity 100.0%; Pred. NO. 0.003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NKGTOQYTDQIE 12
 |||||
 Db 395 NKGTOQYTDQIE 406

RESULT 6
 ABP57267
 ID ABP57267 standard; Protein; 554 AA.
 XX
 XX AC ABP57267;
 XX
 XX 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
 XX
 KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 XX
 PR 19-MAR-2002; 2002US-365956P.
 XX
 PR 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX
 DR WPI; 2003-201510/19.
 XX
 DR N-PSDB; AB259577.
 XX
 PT Detecting a human parvovirus B19 infection in a biological sample to

PT	immunogenic properties, useful in gene therapy protocols -
XX	
PS	Example 21; Page 142; 153pp; English.
XX	
CC	The patent discloses modified parvovirus vectors with advantageous
CC	antigenic properties, packaging capabilities and cellular tropisms.
CC	These vectors can be used in standard recombinant DNA protocols e.g. gene
CC	therapy for delivering nucleic acids to cells.
CC	The present sequence is a capsid protein encoded by an
CC	adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric
CC	vector. This vector encodes AAV2 VP1 and VP2 capsid proteins
CC	and human parvovirus B19 VP2 protein. The chimeric vector was
CC	constructed by replacing the VP3 major cap protein of AAV2 with B19's
CC	VP2. Recombinant parvovirus comprising the chimeric capsid is useful
CC	for gene delivery.
XX	
SQ	Sequence 756 AA;
XX	
Query Match	100.0%; Score 64; DB 21; Length 756;
Best Local Similarity	100.0%; Pred. NO. 0.0043;
Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Qy	1 NKGTQYTDQIE 12
Db	597 NKGTQYTDQIE 608
XX	
RESULT 8	
AAW08986	
ID	AAW08986 standard; Protein; 781 AA.
XX	
AC	AAW08986;
XX	
DT	27-FEB-1997 (first entry)
XX	
DE	Human parvovirus VP-1 protein.
XX	
KM	Human, parvovirus genome; structural gene; VP-1; VP2; arthritis;
KW	non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
KM	erythraemia; abortion; universal fetal hydrops; liver disease;
KM	haemorrhagic fever; rheumatism; detection; IgG antibody.
XX	
OS	Human parvovirus.
XX	
KY	Location/Qualifiers
FT	Misc-difference 19
FT	/label= Gly, Val
FT	Misc-difference 61
FT	/label= Asn, Asp
FT	Misc-difference 220
FT	/label= His, Asn
XX	
PN	JP07147986-A.
XX	
PD	13-JUN-1995.
XX	
PF	24-SEP-1992; 92JP-0281017.
XX	
PR	24-SEP-1992; 92JP-0281017.
XX	
PA	(DENK-) DENKA SEIKEN KK.
PA	(ELED) DENKI KAGAKU KOGYO KK.
XX	
DR	WI: 1995-242756/32.
XX	
N-PSDB	AAT49535.
XX	
PT	Human parvovirus gene coding for a polypeptide - useful for
PT	developing vaccines against parvoviral diseases such as
XX	erythroblastemia, haemorrhagic fever, etc.
XX	
PS	Claim 2; Page 5-7; 38pp; English.
XX	
CC	The sequences given in AAW08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

XX Sequence 781 AA;

Query Match 100.0%; Score 64; DB 16; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.0045;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQIE 12
 |||||
 DB 622 NKGTOOYTDQIE 633

RESULT 9

AAV23227 standard; Protein; 781 AA.

AC AAY23227;

DT 26-AUG-1999 (first entry)

DE Erythrovirus V9 VP1 protein.

XX Erythrovirus V9; differential diagnosis; parvovirus; infection;

KM erythrovirus screening; typing; immunoassay; VP1 protein.

XX Erythrovirus.

PN FR2771751-A1.

PD 04-JUN-1999.

PF 03-DEC-1997; 97FR-0015197.

PR 03-DEC-1997; 97FR-0015197.

PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

PI Auguste V, Garbarg CA, Nguyen QT;

XX WPI, 1999-349543/30.

DR N-PSDB; AAX81583.

XX Erythrovirus V9 and its nucleic acid sequences - can be used in the
 diagnosis of its infections

BS Claim 19; Page 50-52; 80pp; French.

XX The present sequence represents an erythrovirus V9 protein.
 CC probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.

XX Sequence 781 AA;

Query Match 100.0%; Score 64; DB 20; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.0045;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQIE 12
 |||||
 DB 622 NKGTOOYTDQIE 633

RESULT 10

ABP57263 standard; Protein; 781 AA.

XX ABP57263;

DT 22-APR-2003 (first entry)

DE Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

XX Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

OS Human parvovirus B19.

PN WO2003002753-A2.

PD 09-JAN-2003.

PF 28-JUN-2002; 2002WO-US20684.

PR 28-JUN-2001; 2001US-302077P.

PR 19-MAR-2002; 2002US-365956P.

PR 29-MAR-2002; 2002US-369224P.

PA (CHIR) CHIRON CORP.

PI Pichuanes S, Shyamala V;

XX WPI, 2003-201510/19.

DR N-PSDB; ABZ59573.

PT Detecting a human parvovirus B19 infection in a biological sample to
 prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 acid with a primer complementary to the 3'-terminal portion of the RNA
 target sequence

PT Example 4; Fig 6B; 148pp; English.

XX The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59869, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 64; DB 24; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.0045;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOOYTDQIE 12
 |||||
 DB 622 NKGTOOYTDQIE 633

RESULT 11
 ABP57266

ID ABP57266 standard; Protein; 781 AA.
 AC- ABP57266;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B6-VP1 amino acid sequence SEQ ID NO:33.
 XX
 DE Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 KM Human parvovirus B19;
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PR 28-JUN-2001; 2001US-302077P.
 XX 19-MAR-2002; 2002US-365956P.
 XX 29-MAR-2002; 2002US-369224P.
 XX
 PI (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX
 DR WPI; 2003-201510/19.
 XX
 DR N-PSDB; ABZ59576.
 XX
 PT Detecting a human parvovirus B19 infection in a biological sample to
 XX prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 XX acid with a primer complementary to the 3'-terminal portion of the RNA
 XX target sequence -
 XX
 PS Example 4; Fig 9B; 148pp; English.
 XX
 CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 781 AA;
 XX
 Query Match 100.0%; Score 64; DB 24; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.0045;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NKGTOOYTDQIR 12
 |||||
 DB 622 NKGTOOYTDQIR 633
 |||||
 RESULT 12
 ABB60093
 ID ABB60093 standard; Protein; 237 AA.
 XX

AC ABB60093;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7071.
 XX
 DE Drosophila; developmental biology; cell signalling; insecticide;
 XX
 KM Drosophila;
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 DR N-PSDB; ABL04196.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -
 XX
 PS Disclosure; SEQ ID NO 7071; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01640-ABL01675) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 237 AA;
 XX
 Query Match 64.1%; Score 41; DB 22; Length 237;
 Best Local Similarity 70.0%; Pred. No. 19;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTOOYTDQ 10
 |||||
 DB 30 NKGTOOYTDQ 39
 |||||
 RESULT 13
 AAY14929
 ID AAY14929 standard; Protein; 187 AA.
 XX
 AC AAY14929;
 XX
 DT 25-OCT-1999 (first entry)
 XX
 DE Amino acid sequence of M. vaccae antigen GV-41B.
 XX
 KM Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 XX
 KM dendritic cell maturation; infectious disease; immune disorder; cancer;
 XX
 KM respiratory system; mycobacterial infection; allergy; tuberculosis;
 XX
 KM leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
 XX
 KM dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 XX
 KM squamous cell carcinoma; melanoma.
 XX

OS Mycobacterium vaccae.
 XX
 PN WO9932634-A2.
 XX
 PD 01-JUL-1999.
 XX
 PF 23-DEC-1998; 98WO-NZ00189.
 XX
 PR 04-DEC-1998; 98US-0205426.
 PR 23-DEC-1997; 97US-0996624.
 PR 23-DEC-1997; 97US-0997080.
 PR 23-DEC-1997; 97US-0997362.
 PR 11-JUN-1998; 98US-0095855.
 PR 17-SEP-1998; 98US-0156181.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
 XX
 WP1; 1999-430163/36.
 N-PSDB; AAZ11394.
 PF Enhancing immune response to an antigen
 Claim 1; Page 240; 243pp; English.
 XX
 PS The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.
 CC
 SQ Sequence 187 AA;
 XX
 Query Match 62.5%; Score 40; DB 20; Length 187;
 Best Local Similarity 58.3%; Pred. No. 22;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 DB
 1 NKGTQOYTDQIE 12
 :|||:|:|:
 161 DKSTHQYTNQID 172
 RESULT 14
 ABB64125
 ID ABB64125 standard; Protein; 277 AA.
 XX
 AC ABB64125;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 19167.
 XX
 KM Drosophila developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEXE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WO1; 2001-656860/75.
 DR N-PSDB; ABL08228.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PF Disclosure; SEQ ID NO 19167; 21pp + Sequence Listing; English.
 XX
 PS The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 277 AA;
 XX
 Query Match 60.9%; Score 39; DB 22; Length 277;
 Best Local Similarity 60.0%; Pred. No. 53;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB
 1 NKGTQOYTDQ 10
 :|||:|:|:
 247 NRGTRQYDE 256

RESULT 15
 AA015899
 ID AA015899 standard; protein; 522 AA.
 XX
 AC AA015899;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE Pseudomonas glutaryl amidase (GA) enzyme B chain.
 XX
 KM Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA;
 KM 7-amino-cephalosporanic acid.
 XX
 OS Pseudomonas sp.
 XX
 PN WO200272806-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 12-MAR-2002; 2002WO-IB02119.
 XX
 PR 14-MAR-2001; 2001DE-1012608.
 PR 02-OCT-2001; 2001DE-1048723.
 PR 31-OCT-2001; 2001DE-1053389.
 XX
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (KOLL) KOLLER K.
 PA (LANG) LANGE G.
 PA (SAUB) SAUBER K.
 XX
 PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
 XX

DR WPI; 2002-732828/79.

XX New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid

PS Claim 2; Page 28-30; 36pp; English.

CC The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate
CC binding pocket that binds the alpha-amino adityl moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents the B chain of the Pseudomonas
CC glutaryl amidase enzyme.

Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 GTQGYTDOIE 12

DB 482 GTTHYSDOIE 491

RESULT 16

AAOI5900

ID AAOI5900 standard; protein; 522 AA.

AC AAOI5900;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #1.

KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;

OS mutant; 7-amino-cephalosporanic acid.

OS Pseudomonas sp.

OS Synthetic.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid

PS Claim 10; Page -; 36pp; English.

CC The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate
CC binding pocket that binds the alpha-amino adityl moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents a mutant B chain of the
CC Pseudomonas glutaryl amidase enzyme.
CC NOTE: The present sequence is not shown in the specification but is
CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AAOI5899).

Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 GTQGYTDOIE 12

DB 482 GTTHYSDOIE 491

RESULT 17

AAOI5901

ID AAOI5901 standard; protein; 522 AA.

AC AAOI5901;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #2.

KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;

OS mutant; 7-amino-cephalosporanic acid.

OS Pseudomonas sp.

OS Synthetic.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Key Location/Qualifiers

Misc-difference 24 /note= "Wild type Leu replaced by Arg"

Misc-difference 57 /note= "Wild type Arg replaced by His"

Misc-difference 177 /note= "Wild type Gln replaced by Val"

Misc-difference 57 /note= "Wild type Arg replaced by His"

Misc-difference 177 /note= "Wild type Phe replaced by His"

WO200272806-A2.

19-SEP-2002.

12-MAR-2002; 2002MO-IB02119.

14-MAR-2001; 2001DE-1012608.

02-OCT-2001; 2001DE-1048723.

31-OCT-2001; 2001DE-1053389.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

(KOELL/) KOLLER K.

(LANG/) LANGE G.

(SAUB/) SAUBER K.

(SAUB/) SAUBER K.

PR 31-OCT-2001; 2001DE-1053389
PR

PD 19-SEP-2002

PF 12-MAR-2002; 2002WO-1B02119.
 XX
 XX 14-MAR-2001; 2001DE-1012608.
 PR 02-OCT-2001; 2001DE-1048723.
 PR 31-OCT-2001; 2001DE-1053389.
 XX
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (KOLL/) KOLLER K.
 PA (LANG/) LANGE G.
 PA (SAUB/) SAUBER K.
 XX
 PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W,
 XX
 DR WPI; 2002-732828/79.
 XX
 PT New non-naturally occurring variant of glutaryl amidase comprising
 PT histidine or glutamate in its substrate binding pocket, binds
 PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 PT 7-aminocephalosporanic acid

Claim 10; Page -; 36pp; English.

CC The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AA015899).

SQ Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
 Best Local Similarity 70.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTQOYTDQIE 12
 |||:||||
 DB 482 GTTHYSDQIE 491

RESULT 20

5904
 AA015904 standard; protein; 522 AA.

XX AA015904;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #5.

XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KM mutetin; 7-amino-cephalosporanic acid.

XX Pseudomonas sp.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 24 /note= "Wild type Leu replaced by Arg"

FT Misc-difference 50 /note= "Wild type Gln replaced by Ala"

FT Misc-difference 57 /note= "Wild type Arg replaced by His"

FT Misc-difference 70 /note= "Wild type Val replaced by His"

FT Misc-difference 177

/note= "Wild type Phe replaced by His"

PN W0200272806-A2.

PD 19-SEP-2002.

PF 12-MAR-2002; 2002WO-1B02119.

PR 14-MAR-2001; 2001DE-1012608.

PR 02-OCT-2001; 2001DE-1048723.

PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX (KOLL/) KOLLER K.

XX (LANG/) LANGE G.

XX (SAUB/) SAUBER K.

XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;
 XX
 DR WPI; 2002-732828/79.

PT New non-naturally occurring variant of glutaryl amidase comprising
 PT histidine or glutamate in its substrate binding pocket, binds
 PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 PT 7-aminocephalosporanic acid

Claim 10; Page -; 36pp; English.

CC The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AA015899).

SQ Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
 Best Local Similarity 70.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTQOYTDQIE 12
 |||:||||
 DB 482 GTTHYSDQIE 491

RESULT 21

AA015905
 AA015905 standard; protein; 522 AA.

XX AA015905;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #6.

XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KM mutetin; 7-amino-cephalosporanic acid.

XX Pseudomonas sp.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 24 /note= "Wild type Leu replaced by Arg"

FT Misc-difference 50


```

AC  AAO15907;
XX
XX  30-JAN-2003 (first entry)
XX
DE  Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #8.
XX
XX  Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX  muten; 7-amino-cephalosporanic acid.
XX
OS  Pseudomonas sp.
OS  Synthetic.
XX
XX  Key Location/Qualifiers
XX  FH MISC-difference 176 /note= "Wild type Thr replaced by Asp"
XX  FT MISC-difference 177 /note= "Wild type Phe replaced by His"
XX  FT MISC-difference 177 /note= "Wild type Phe replaced by His"
XX
XX  WO200272806-A2.
XX
XX  19-SEP-2002.
XX
XX  12-MAR-2002; 2002WO-IB02119.
XX
XX  14-MAR-2001; 2001DE-1012608.
XX  02-OCT-2001; 2001DE-1048723.
XX  31-OCT-2001; 2001DE-1053389.
XX
XX  (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX  (KOLL/) KOLLER K.
XX  (LANG/) LANGE G.
XX  (SAUB/) SAUBER K.
XX
XX  Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
XX
XX  WPI; 2002-732828/79.
XX
XX  New non-naturally occurring variant of glutaryl amidase comprising
XX  histidine or glutamate in its substrate binding pocket, binds
XX  cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX  7-aminocephalosporanic acid -
XX
XX  Claim 10; Page -; 36pp; English.
XX
XX  The invention comprises the amino acid sequences of mutant Pseudomonas
XX  glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
XX  bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
XX  CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
XX  the invention comprise a histidine or glutamate residue in the substrate
XX  binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
XX  mutant GA proteins of the invention are useful for preparing 7-ACA from
XX  CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
XX  present amino acid sequence represents a mutant B chain of the
XX  Pseudomonas glutaryl amidase enzyme.
XX  NOTE: The present sequence is not shown in the specification but is
XX  derived from the Pseudomonas wild type glutaryl amidase B chain, (see
XX  AAO15899).
XX
XX  Sequence 522 AA;
XX
XX  Query Match 59.4%; Score 38; DB 23; Length 522;
XX  Best Local Similarity 70.0%; Pred. No. 1.7e+02;
XX  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY  3 GTGQYTDQIE 12
XX  |||:|||||
XX  482 GTTHYSDDIE 491
XX
RESULT 24
AAO15908
ID  AAO15908 standard; protein; 522 AA.
XX

```

```

AC  AAO15908;
XX
XX  30-JAN-2003 (first entry)
XX
XX  Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #9.
XX
XX  Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX  muten; 7-amino-cephalosporanic acid.
XX
OS  Pseudomonas sp.
OS  Synthetic.
XX
XX  Key Location/Qualifiers
XX  FH MISC-difference 24 /note= "Wild type Leu replaced by Arg"
XX  FT MISC-difference 57 /note= "Wild type Arg replaced by His"
XX  FT MISC-difference 176 /note= "Wild type Thr replaced by Asp"
XX  FT MISC-difference 177 /note= "Wild type Phe replaced by His"
XX  FT MISC-difference 177 /note= "Wild type Phe replaced by His"
XX
XX  WO200272806-A2.
XX
XX  19-SEP-2002.
XX
XX  12-MAR-2002; 2002WO-IB02119.
XX
XX  14-MAR-2001; 2001DE-1012608.
XX  02-OCT-2001; 2001DE-1048723.
XX  31-OCT-2001; 2001DE-1053389.
XX
XX  (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX  (KOLL/) KOLLER K.
XX  (LANG/) LANGE G.
XX  (SAUB/) SAUBER K.
XX
XX  Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
XX
XX  WPI; 2002-732828/79.
XX
XX  New non-naturally occurring variant of glutaryl amidase comprising
XX  histidine or glutamate in its substrate binding pocket, binds
XX  cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX  7-aminocephalosporanic acid -
XX
XX  Claim 10; Page -; 36pp; English.
XX
XX  The invention comprises the amino acid sequences of mutant Pseudomonas
XX  glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
XX  bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
XX  CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
XX  the invention comprise a histidine or glutamate residue in the substrate
XX  binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
XX  mutant GA proteins of the invention are useful for preparing 7-ACA from
XX  CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
XX  present amino acid sequence represents a mutant B chain of the
XX  Pseudomonas glutaryl amidase enzyme.
XX  NOTE: The present sequence is not shown in the specification but is
XX  derived from the Pseudomonas wild type glutaryl amidase B chain, (see
XX  AAO15899).
XX
XX  Sequence 522 AA;
XX
XX  Query Match 59.4%; Score 38; DB 23; Length 522;
XX  Best Local Similarity 70.0%; Pred. No. 1.7e+02;
XX  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY  3 GTGQYTDQIE 12
XX  |||:|||||
XX  482 GTTHYSDDIE 491
XX

```

RESULT 25
 AA015910
 ID ID AA015910 standard; protein. 522 AA.
 XX AC AA015910;
 XX AC AA015910;
 DT 30-JAN-2003 (first entry)
 XX
 DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #10.
 XX
 XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 KM muten; 7-amino-cephalosporanic acid.
 XX
 OS Pseudomonas sp.
 OS Synthetic.
 XX
 XX
 FH Location/Qualifiers
 FT Key 50
 FT Misc-difference /note= "Wild type Gln replaced by Arg"
 FT Misc-difference 57 /note= "Wild type Arg replaced by His"
 FT Misc-difference 176 /note= "Wild type Thr replaced by Asp"
 FT Misc-difference 177 /note= "Wild type Phe replaced by His"
 XX
 XX WO200272806-A2.
 XX
 XX 19-SEP-2002.
 PD 12-MAR-2002; 2002WO-IB02119.
 XX
 XX 14-MAR-2001; 2001DE-1012608.
 PR 02-OCT-2001; 2001DE-1048723.
 PR 31-OCT-2001; 2001DE-1053389.
 XX
 XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (KOLL/) KOLLER K.
 PA (LANG/) LANG G.
 PA (SAUB/) SAUBER K.
 XX
 PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
 XX
 XX WPI; 2002-732828/79.
 DR
 XX
 XX
 PT New non-naturally occurring variant of glutaryl amidase comprising
 PT histidine or glutamate in its substrate binding pocket, binds
 PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 7-aminocephalosporanic acid -
 PS
 PS Claim 10; Page -; 36pp; English.
 CC
 CC The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AA015899).
 CC
 XX
 SO Sequence 522 AA;
 Query Match 59.4%; Score 38; DB 23; Length 522;
 Best Local Similarity 70.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0

```

Db          482 GTTHYSDQIE 491

RESULT 26
AA015913
ID .AA015913 standard; protein; 522 AA.
XX
AC
XX
AA015913;
XX
30-JAN-2003 (first entry)
XX
DE
Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #11.
XX
KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
KW muten; 7-amino-cephalosporanic acid.
XX
OS Pseudomonas sp.
XX Synthetic.
XX
Key Location/Qualifiers
FH Misc-difference 177
FT /note= "Wild type Phe replaced by His"
XX
XX MO200272806-A2.
XX
XX 19-SEP-2002.
XX
XX 12-MAR-2002; 2002MO-IB02119.
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC ) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.
XX (KOLL/) KOLLER K.
XX (LANG/) LANGE G.
XX (SAUB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabsch W;
XX
XX WPI; 2002-732826/79.
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
XX histidine or glutamate in its substrate binding pocket, binds
XX cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.
XX
XX The invention comprises the amino acid sequences of mutant Pseudomonas
XX glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
XX bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
XX CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
XX the invention comprise a histidine or glutamate residue in the substrate
XX binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
XX mutant GA proteins of the invention are useful for preparing 7-ACA from
XX CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
XX present amino acid sequence represents a mutant B chain of the
XX Pseudomonas glutaryl amidase enzyme.
XX NOTE: The present sequence is not shown in the specification but is
XX derived from the Pseudomonas wild type glutaryl amidase B chain, (see
XX AA015899).
XX
XX Sequence 522 AA;
XX
XX Query Match 59.4%; Score 38; DB 23; Length 522;
XX Best Local Similarity 70.0%; Pred. No. 1.7e+02;
XX Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0.
XX
XX 3 GTGQYTDQIE 12
XX |||:|||||
XX 482 GTTHYSDQIE 491

```

RESULT 27

AAO15914 standard; protein; 522 AA.

AAO15914;

30-JAN-2003 (first entry)

Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #12.

Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;

mutatein; 7-amino-cephalosporanic acid.

Pseudomonas sp.

Synthetic.

Location/Qualifiers

Key Misc-difference 24 /note= "Wild type Leu replaced by Arg"

Misc-difference 57 /note= "Wild type Arg replaced by His"

WO200272806-A2.

19-SEP-2002.

12-MAR-2002; 2002WO-1B02119.

14-MAR-2001; 2001DE-1012608.

02-OCT-2001; 2001DE-1048723.

31-OCT-2001; 2001DE-1053389.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

(KOLL/) KOLLER K.

(LANG/) LANGE G.

(SAUB/) SAUBER K.

Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;

WPI; 2002-732828/79.

New non-naturally occurring variant of glutaryl amidase comprising

histidine or glutamate in its substrate binding pocket, binds

cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to

7-aminocephalosporanic acid

Claim 10; Page -; 36pp; English.

The invention comprises the amino acid sequences of mutant Pseudomonas

glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention

bind cephalosporin C (CPC) as a substrate and catalyze the conversion of

CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of

the invention comprise a histidine or glutamate residue in the substrate

binding pocket that binds the alpha-amino adipyl moiety of the CPC. The

mutant GA proteins of the invention are useful for preparing 7-ACA. The

CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The

present amino acid sequence represents a mutant B chain of the

Pseudomonas glutaryl amidase enzyme.

NOTE: The present sequence is not shown in the specification but is

derived from the Pseudomonas wild type glutaryl amidase B chain, (see

AAO15899).

Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;

Best Local Similarity 70.0%; Pred. No. 1.7e+02; 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 2; Indels 0;

3 GTGQYTDQIE 12

482 GTTHYSQDIE 491

RESULT 28

AAO15915 standard; protein; 522 AA.

AAO15915;

30-JAN-2003 (first entry)

Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #13.

Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;

mutatein; 7-amino-cephalosporanic acid.

Pseudomonas sp.

Synthetic.

Location/Qualifiers

Key Misc-difference 24 /note= "Wild type Leu replaced by Arg"

Misc-difference 57 /note= "Wild type Arg replaced by His"

Misc-difference 177 /note= "Wild type Phe replaced by His"

WO200272806-A2.

19-SEP-2002.

12-MAR-2002; 2002WO-1B02119.

14-MAR-2001; 2001DE-1012608.

02-OCT-2001; 2001DE-1048723.

31-OCT-2001; 2001DE-1053389.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

(KOLL/) KOLLER K.

(LANG/) LANGE G.

(SAUB/) SAUBER K.

Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;

WPI; 2002-732828/79.

New non-naturally occurring variant of glutaryl amidase comprising

histidine or glutamate in its substrate binding pocket, binds

cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to

7-aminocephalosporanic acid

Claim 10; Page -; 36pp; English.

The invention comprises the amino acid sequences of mutant Pseudomonas

glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention

bind cephalosporin C (CPC) as a substrate and catalyze the conversion of

CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of

the invention comprise a histidine or glutamate residue in the substrate

binding pocket that binds the alpha-amino adipyl moiety of the CPC. The

mutant GA proteins of the invention are useful for preparing 7-ACA. The

CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The

present amino acid sequence represents a mutant B chain of the

Pseudomonas glutaryl amidase enzyme.

NOTE: The present sequence is not shown in the specification but is

derived from the Pseudomonas wild type glutaryl amidase B chain, (see

AAO15899).

Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;

Best Local Similarity 70.0%; Pred. No. 1.7e+02; 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 2; Indels 0;

3 GTGQYTDQIE 12

Db 482 GTTHYSQDIE 491

RESULT 29

AA015916
ID AA015916 standard; protein; 522 AA.
XX
AC AA015916;
XX
DT 30-JAN-2003 (first entry)
XX
DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #14.
XX
KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX murein; 7-amino-cephalosporanic acid.
XX
OS Pseudomonas sp.
OS Synthetic.

Key Location/Qualifiers

FT Misc-difference 50 /note= "Wild type Gln replaced by Arg"
FT Misc-difference 57 /note= "Wild type Arg replaced by His"
XX
XX WO200272806-A2.
XX
XX 19-SEP-2002.
XX
XX 12-MAR-2002; 2002WO-1B02119.
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (KOLL/) KOLLER K.
XX (LANG/) LANGE G.
XX (SAUB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;
XX
XX WPI; 2002-732828/79.
XX
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
XX histidine or glutamate in its substrate binding pocket, binds
XX cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.

The invention comprises the amino acid sequences of mutant Pseudomonas
glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
the invention comprise a histidine or glutamate residue in the substrate
binding pocket that binds the alpha-amino adityl moiety of the CPC. The
mutant GA proteins of the invention are useful for preparing 7-ACA from
CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
present amino acid sequence represents a mutant B chain of the
Pseudomonas glutaryl amidase enzyme.
NOTE: The present sequence is not shown in the specification but is
derived from the Pseudomonas wild type glutaryl amidase B chain, (see
AA015899).

Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 GTQOYTDI012

Db 482 GTTHYSQDIE 491

RESULT 30

AA015917
ID AA015917 standard; protein; 522 AA.
XX
AC AA015917;
XX
DT 30-JAN-2003 (first entry)
XX
DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #15.
XX
KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX murein; 7-amino-cephalosporanic acid.
XX
OS Pseudomonas sp.
OS Synthetic.

Key Location/Qualifiers

FT Misc-difference 50 /note= "Wild type Gln replaced by Arg"
FT Misc-difference 57 /note= "Wild type Arg replaced by His"
XX
XX WO200272806-A2.
XX
XX 19-SEP-2002.
XX
XX 12-MAR-2002; 2002WO-1B02119.
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (KOLL/) KOLLER K.
XX (LANG/) LANGE G.
XX (SAUB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;
XX
XX WPI; 2002-732828/79.
XX
XX
XX New non-naturally occurring variant of glutaryl amidase comprising
XX histidine or glutamate in its substrate binding pocket, binds
XX cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
XX 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.

The invention comprises the amino acid sequences of mutant Pseudomonas
glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
the invention comprise a histidine or glutamate residue in the substrate
binding pocket that binds the alpha-amino adityl moiety of the CPC. The
mutant GA proteins of the invention are useful for preparing 7-ACA from
CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
present amino acid sequence represents a mutant B chain of the
Pseudomonas glutaryl amidase enzyme.
NOTE: The present sequence is not shown in the specification but is
derived from the Pseudomonas wild type glutaryl amidase B chain, (see
AA015899).

Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 GTQYTDQIE 12
 ||:||||
 DB 482 GTTHYSDQIE 491

RESULT 31

AAO15918
 ID AAO15918 standard; protein; 522 AA.

AC AAO15918;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #16.

KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 mutant; 7-amino-cephalosporanic acid.

XX Pseudomonas sp.
 Synthetic.

AM Key Location/Qualifiers

FT Misc-difference 33 /note= "Wild type Tyr replaced by Ser"

FT Misc-difference 176 /note= "Wild type Thr replaced by Asp"

FT Misc-difference 177 /note= "Wild type Phe replaced by His"

PN WO200272806-A2.

PD 19-SEP-2002.

PF 12-MAR-2002; 2002WO-IB02119.

PR 14-MAR-2001; 2001DE-1012608.

PR 02-OCT-2001; 2001DE-1048723.

PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (KOLL/) KOLLER K.

PA (LANG/) LANGE G.

PA (SAUB/) SAUBER K.

PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;

PI WPI; 2002-732828/79.

PS Claim 10; Page -; 36pp; English.

XX The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.

CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AAO15899).

XX Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;

Best Local Similarity 70.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 GTQYTDQIE 12
 ||:||||
 DB 482 GTTHYSDQIE 491

RESULT 32

AAO15919
 ID AAO15919 standard; protein; 522 AA.

AC AAO15919;

DT 30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #17.

KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 mutant; 7-amino-cephalosporanic acid.

XX Pseudomonas sp.
 Synthetic.

OS Key Location/Qualifiers

FT Misc-difference 24 /note= "Wild type Leu replaced by Arg"

FT Misc-difference 33 /note= "Wild type Tyr replaced by Ser"

FT Misc-difference 57 /note= "Wild type Arg replaced by His"

FT Misc-difference 176 /note= "Wild type Thr replaced by Asp"

FT Misc-difference 177 /note= "Wild type Phe replaced by His"

PN WO200272806-A2.

PD 19-SEP-2002.

PF 12-MAR-2002; 2002WO-IB02119.

PR 14-MAR-2001; 2001DE-1012608.

PR 02-OCT-2001; 2001DE-1048723.

PR 31-OCT-2001; 2001DE-1053389.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (KOLL/) KOLLER K.

PA (LANG/) LANGE G.

PA (SAUB/) SAUBER K.

PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;

PI WPI; 2002-732828/79.

PS Claim 10; Page -; 36pp; English.

XX The invention comprises the amino acid sequences of mutant Pseudomonas
 CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
 CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.

CC NOTE: The present sequence is not shown in the specification but is

CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AAO15899).

XX
SQ Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTQGYTDQIE 12
|||:||||
Db 482 GTTHYSDQIE 491

RESULT 33
AAO15920
ID AAO15920 standard; protein; 522 AA.
XX
AC AAO15920;

30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #18.
XX
KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
KW mutcin; 7-amino-cephalosporanic acid.
XX
OS Pseudomonas sp.
OS Synthetic.

XX
FH Key Location/Qualifiers
FH Misc-difference 33 /note= "Wild type Tyr replaced by Ser"
FT Misc-difference 50 /note= "Wild type Gln replaced by Arg"
FT Misc-difference 57 /note= "Wild type Arg replaced by His"
FT Misc-difference 176 /note= "Wild type Thr replaced by Asp"
FT Misc-difference 177 /note= "Wild type Phe replaced by His"
FT
XX
PN WO200272806-A2.
XX
PD 19-SEP-2002.
XX
PF 12-MAR-2002; 2002WO-IB02119.
XX
PR 14-MAR-2001; 2001DE-1012608.
PR 02-OCT-2001; 2001DE-1048723.
PR 31-OCT-2001; 2001DE-1053389.

XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (KOLL/) KOLLER K.
PA (LANG/) LANGE G.
PA (SAUB/) SAUBER K.
XX
PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
XX
DR WPI; 2002-732828/79.

XX
PT New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
PT 7-aminocephalosporanic acid -
XX
PS Claim 10; Page -; 36pp; English.

XX
CC The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate

CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents a mutant B chain of the
CC Pseudomonas glutaryl amidase enzyme.
CC NOTE: The present sequence is not shown in the specification but is
CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AAO15899).

XX
SQ Sequence 522 AA;

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTQGYTDQIE 12
|||:||||
Db 482 GTTHYSDQIE 491

RESULT 34
AAO15921
ID AAO15921 standard; protein; 522 AA.
XX
AC AAO15921;

30-JAN-2003 (first entry)

DE Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #19.
XX
KW Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
KW mutcin; 7-amino-cephalosporanic acid.
XX
OS Pseudomonas sp.
OS Synthetic.

XX
FH Key Location/Qualifiers
FH Misc-difference 33 /note= "Wild type Tyr replaced by Ser"
FT Misc-difference 177 /note= "Wild type Phe replaced by His"
FT
XX
PN WO200272806-A2.
XX
PD 19-SEP-2002.
XX
PF 12-MAR-2002; 2002WO-IB02119.
XX
PR 14-MAR-2001; 2001DE-1012608.
PR 02-OCT-2001; 2001DE-1048723.
PR 31-OCT-2001; 2001DE-1053389.

XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (KOLL/) KOLLER K.
PA (LANG/) LANGE G.
PA (SAUB/) SAUBER K.
XX
PI Koller K, Lange G, Sauber K, Fritz-wolf K, Kabesch W;
XX
DR WPI; 2002-732828/79.

XX
PT New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
PT 7-aminocephalosporanic acid -
XX
PS Claim 10; Page -; 36pp; English.

XX
CC The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyse the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate

CC binding pocket that binds the alpha-amino adipyrl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AAO15899).

CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
 CC the invention comprise a histidine or glutamate residue in the substrate
 CC binding pocket that binds the alpha-amino adipyrl moiety of the CPC. The
 CC mutant GA proteins of the invention are useful for preparing 7-ACA from
 CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
 CC present amino acid sequence represents a mutant B chain of the
 CC Pseudomonas glutaryl amidase enzyme.
 CC NOTE: The present sequence is not shown in the specification but is
 CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
 CC AAO15899).

Query Match 59.4%; Score 38; DB 23; Length 522;
 Best Local Similarity 70.0%; Pred. NO. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Sequence 522 AA;

QY 3 GTQGYTDQIE 12
 |||||
 DB 482 GTTHYSDQIE 491

QY 3 GTQGYTDQIE 12
 |||||
 DB 482 GTTHYSDQIE 491

AAO15922 standard; protein; 522 AA.

AAO15922;

30-JAN-2003 (first entry)

Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #20.

glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 mutetrn; 7-amino-cephalosporanic acid.

Pseudomonas sp.
 Synthetic.

Location/Qualifiers

Key

Misc-difference 24

/note= "Wild type Leu replaced by Arg"

Misc-difference 33

/note= "Wild type Tyr replaced by Ser"

Misc-difference 57

/note= "Wild type Arg replaced by His"

WO200272806-A2.

19-SEP-2002.

12-MAR-2002; 2002WO-IB02119.

14-MAR-2001; 2001DE-1012608.

02-OCT-2001; 2001DE-1048723.

31-OCT-2001; 2001DE-1053389.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

(KOLL/) KOLLER K.

(LANG/) LANGE G.

(SAUB/) SAUBER K.

Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;

WPI; 2002-732828/79.

New non-naturally occurring variant of glutaryl amidase comprising
 histidine or glutamate in its substrate binding pocket, binds
 cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 7-aminocephalosporanic acid -
 Claim 10; Page -; 36pp; English.

The invention comprises the amino acid sequences of mutant Pseudomonas
 glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
 bind cephalosporin C (CPC) as a substrate and catalyse the conversion of

RESULT 36

AAO15923 standard; protein; 522 AA.

AAO15923;

30-JAN-2003 (first entry)

Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #21.

glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
 mutetrn; 7-amino-cephalosporanic acid.

Pseudomonas sp.
 Synthetic.

Location/Qualifiers

Key

Misc-difference 24

/note= "Wild type Leu replaced by Arg"

Misc-difference 33

/note= "Wild type Tyr replaced by Ser"

Misc-difference 57

/note= "Wild type Arg replaced by His"

Misc-difference 177

/note= "Wild type Phe replaced by His"

WO200272806-A2.

19-SEP-2002.

12-MAR-2002; 2002WO-IB02119.

14-MAR-2001; 2001DE-1012608.

02-OCT-2001; 2001DE-1048723.

31-OCT-2001; 2001DE-1053389.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

(KOLL/) KOLLER K.

(LANG/) LANGE G.

(SAUB/) SAUBER K.

Koller K, Lange G, Sauber K, Fritz-wolf K, Kabach W;

WPI; 2002-732828/79.

New non-naturally occurring variant of glutaryl amidase comprising
 histidine or glutamate in its substrate binding pocket, binds
 cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
 7-aminocephalosporanic acid -
 Claim 10; Page -; 36pp; English.

XX The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate
CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents a mutant B chain of the
CC Pseudomonas glutaryl amidase enzyme.
CC NOTE: The present sequence is not shown in the specification but is
CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AA015899).

XX Sequence 522 AA;
SQ

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 GTQYTDQIE 12
|||:||||
482 GTTHYSDQIE 491

Db

RESULT 37
AA015924
ID AA015924 standard; protein; 522 AA.
XX AA015924;
XX
XX
XX 30-JAN-2003 (first entry)
XX
XX Mutant Pseudomonas glutaryl amidase (GA) enzyme B chain #22.
XX
XX Glutaryl amidase; GA; enzyme; cephalosporin C; CPC; 7-ACA; mutant;
XX mutacin; 7-amino-cephalosporanic acid.
XX
XX Pseudomonas sp.
XX Synthetic.
XX
XX
XX Key Location/Qualifiers
FH Misc-difference 24
FT /note= "Wild type Leu replaced by Arg"
FT
FT Misc-difference 33 /note= "Wild type Tyr replaced by Ser"
FT
FT Misc-difference 57 /note= "Wild type Arg replaced by His"
FT
FT Misc-difference 70 /note= "Wild type Val replaced by His"
FT
FT Misc-difference 177 /note= "Wild type Phe replaced by His"
FT
FT
FT WO200272806-A2.
XX
XX
XX 19-SEP-2002.
XX
XX
XX 12-MAR-2002; 2002MO-IB02119.
XX
XX
XX 14-MAR-2001; 2001DE-1012608.
XX 02-OCT-2001; 2001DE-1048723.
XX 31-OCT-2001; 2001DE-1053389.
XX
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (KOLL/) KOLLER K.
XX (LANG/) LANGER G.
XX (SAUB/) SAUBER K.
XX
XX Koller K, Lange G, Sauber K, Fritzwolf K, Kabesch W,
XX
XX WPI; 2002-732828/79.
XX

PT New non-naturally occurring variant of glutaryl amidase comprising
PT histidine or glutamate in its substrate binding pocket, binds
PT cephalosporin C (CPC) as substrate and catalyzes conversion of CPC to
PT 7-aminocephalosporanic acid -
XX
XX Claim 10; Page -; 36pp; English.

XX
XX
XX The invention comprises the amino acid sequences of mutant Pseudomonas
CC glutaryl amidase (GA) enzymes. The mutant GA proteins of the invention
CC bind cephalosporin C (CPC) as a substrate and catalyze the conversion of
CC CPC to 7-amino-cephalosporanic acid (7-ACA). The mutant GA proteins of
CC the invention comprise a histidine or glutamate residue in the substrate
CC binding pocket that binds the alpha-amino adipyl moiety of the CPC. The
CC mutant GA proteins of the invention are useful for preparing 7-ACA from
CC CPC - the mutant GA enzyme cleaves the CPC directly to form 7-ACA. The
CC present amino acid sequence represents a mutant B chain of the
CC Pseudomonas glutaryl amidase enzyme.
CC NOTE: The present sequence is not shown in the specification but is
CC derived from the Pseudomonas wild type glutaryl amidase B chain, (see
CC AA015899).

XX Sequence 522 AA;
SQ

Query Match 59.4%; Score 38; DB 23; Length 522;
Best Local Similarity 70.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3 GTQYTDQIE 12
|||:||||
482 GTTHYSDQIE 491

Db

RESULT 38
AAB97344
ID AAB97344 standard; Protein; 691 AA.
XX
XX AAB97344;
XX
XX
XX 14-AUG-2001 (first entry)
XX
XX
XX GL7ACA (acylase) amino acid sequence.
XX
XX
XX GL-7ACA; glutaryl-7-aminocephalophytenic acid acylase.
XX
XX
XX Pseudomonas sp.
XX
XX
XX Key Location/Qualifiers
FH Misc-difference 127
FT /note= "Encoded by TGA"
FT
FT Misc-difference 128 /note= "Encoded by CTG"
FT
FT Misc-difference 549 /note= "Encoded by ACG"
FT
FT Misc-difference 550 /note= "Encoded by GCG"
FT
FT Misc-difference 551 /note= "Encoded by GCG"
FT
FT Misc-difference 552 /note= "Encoded by TCC"
FT
FT Misc-difference 552 /note= "Encoded by GCT"
FT
FT
FT CN1283697-A.
XX
XX
XX 14-FEB-2001.
XX
XX
XX 08-SEP-2000; 2000CN-0125106.
XX
XX 08-SEP-2000; 2000CN-0125106.
XX
XX (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
XX Wang E, Li Y, Zhang Y,
XX
XX WPI; 2001-300983/32.
XX

DR N-PSDB; AAH27731.
 XX Secretion-type bacterial strain with chloromycetin resistance to
 PT generate glutaryl-7-aminocephalophytic acid acylase with high output -
 XX
 PS Claim 1; Page 2-4; 21pp; Chinese.
 CC This invention relates to a recombinant GL-7ACA
 CC (glutaryl-7-aminocephalophytic acid acylase) gene. The gene is obtained
 CC from Pseudomonas and modified using genetic engineering techniques,
 CC involving an expression element and signal peptide. Included in the
 CC invention is a secretion-type bacterial strain with chloromycetin
 CC resistance which generates GL-7ACA at high levels. The present sequence
 CC represents the mature GL-7ACA protein of the invention.
 XX
 SQ Sequence 691 AA;
 Query Match 59.4%; Score 38; DB 22; Length 691;
 Best Local Similarity 70.0%; Pred. No. 2.3e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 3 GTQOYTDQIE 12
 |||:|||||
 651 GTTHYSQDIE.660
 Db
 RESULT 39
 AAB97346
 ID AAB97346 standard; Protein; 691 AA.
 XX
 AC AAB97346;
 XX
 DT 14-AUG-2001 (first entry)
 XX
 DE Mature GL-7ACA amino acid sequence.
 XX
 KW Glutaryl-7-aminocephalophytic acid acylase; GL-7ACA.
 XX
 OS Pseudomonas sp.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 127 /note= "Encoded by TGA"
 FT Misc-difference 128 /note= "Encoded by CTG"
 FT Misc-difference 549 /note= "Encoded by ACG"
 FT Misc-difference 550 /note= "Encoded by GCG"
 FT Misc-difference 551 /note= "Encoded by TCC"
 FT Misc-difference 552 /note= "Encoded by GCT"
 FT /note= "Encoded by GCT"
 PN CN1283696-A.
 PD 14-FEB-2001.
 PF 08-SEP-2000; 2000CN-0125105.
 PR 08-SEP-2000; 2000CN-0125105.
 XX
 PA (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
 XX
 PI Wang E, Zheng Y, Chen J;
 XX
 DR WPI; 2001-291758/31.
 DR N-PSDB; AAH27739.
 XX
 PT Secretion-type bacterial strain to generate
 PT glutaryl-7-aminocephalophytic acid acylase with high output -
 XX
 PS Claim 1; Page 2-3; 20pp; Chinese.

XX
 CC This invention relates to a recombinant (glutaryl-7-aminocephalophytic
 CC acid acylase) GL-7ACA gene. The gene is obtained from pseudomonas and
 CC modified using genetic engineering techniques. The invention includes a
 CC GL-7ACA expression element and signal peptide. A secretion-type bacterial
 CC strain that is able to generate GL-7ACA at high levels is included in the
 CC invention. The present sequence represents mature GL-7ACA of the
 CC invention.
 XX
 SQ Sequence 692 AA;
 Query Match 59.4%; Score 38; DB 22; Length 692;
 Best Local Similarity 70.0%; Pred. No. 2.3e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 3 GTQOYTDQIE 12
 |||:|||||
 652 GTTHYSQDIE 661
 Db
 RESULT 40
 AAG64422
 ID AAG64422 standard; Protein; 692 AA.
 XX
 AC AAG64422;
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE GL-7ACA acylase.
 XX
 KW Heteromer peptide; chitin cellulose binding domain; CBD;
 KW protein immobilisation.
 XX
 OS Unidentified.
 XX
 PN WO200131038-A1.
 XX
 PD 03-MAY-2001.
 XX
 PF 19-OCT-2000; 2000WO-JP07275.
 XX
 PR 22-OCT-1999; 99JP-0301699.
 XX
 PA (FUJI) FUJISAWA PHARM CO LTD.
 XX
 PI Tanaka A, Ueda M, Nagao K;
 XX
 DR WPI; 2001-417561/44.
 DR N-PSDB; AAH46094.
 XX
 PT Heteromer peptides produced by auto-splicing of a precursor comprise
 PT subunits with chitin cellulose binding domain -
 XX
 PS Example 1; Page 61-63; 100pp; Japanese.
 XX
 CC The present sequence is that of the GL-7ACA acylase of the
 CC invention. The invention relates to heteromer peptides, produced by
 CC cleaving precursor peptides, comprise at least one subunit with an added
 CC chitin cellulose binding domain (CBD) useful for immobilising proteins
 CC for industrial use. The CBD fusion allows immobilisation of proteins
 CC without loss of activity and immobilised heteromer peptides may be
 CC produced with high yield.
 XX
 SQ Sequence 692 AA;
 Query Match 59.4%; Score 38; DB 22; Length 692;
 Best Local Similarity 70.0%; Pred. No. 2.3e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 3 GTQOYTDQIE 12
 |||:|||||
 652 GTTHYSQDIE 661
 Db

Search completed: August 20, 2003, 09:29:35
Job time : 40.8247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 13.2371 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-4
Perfect score: 64
Sequence: 1 NKG2QXYTDQIE 12

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_Aa.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	395	3	US-08-856-841-13
2	64	100.0	398	3	US-08-856-841-21
3	64	100.0	415	3	US-08-856-841-20
4	64	100.0	543	3	US-08-856-841-22
5	64	100.0	756	4	US-09-438-268-4
6	40	62.5	187	3	US-09-095-855-203
7	40	62.5	187	4	US-07-205-426-203
8	38	59.4	720	2	US-08-541-760-2
9	38	57.8	373	3	US-09-039-198A-14
10	37	57.8	373	3	US-09-039-198A-15
11	37	57.8	373	4	US-08-877-599-14
12	37	57.8	373	4	US-08-877-599-15
13	37	57.8	373	4	US-09-267-574-14
14	37	57.8	373	4	US-09-267-574-15
15	37	57.8	387	2	US-08-486-839-6
16	37	57.8	387	2	US-09-151-011-6
17	37	57.8	387	4	US-09-343-623-6
18	37	57.8	387	4	US-09-343-623-7
19	37	57.8	415	4	US-09-554-999-2
20	37	57.8	466	2	US-08-486-839-4
21	37	57.8	466	3	US-09-151-011-4
22	37	57.8	466	3	US-09-039-198A-2
23	37	57.8	466	4	US-09-343-623-4
24	37	57.8	466	4	US-08-877-599-2
25	37	57.8	466	4	US-09-267-574-2
26	36	56.2	738	4	US-09-134-001C-4285
27	36	56.2	863	3	US-09-238-303-11

28	36	56.2	863	4	US-09-946-239-11	Sequence 11, Appl
29	35	54.7	61	4	US-09-328-352-7978	Sequence 7978, Ap
30	35	54.7	91	4	US-09-198-452A-753	Sequence 753, App
31	34	53.1	137	3	US-09-456-830-7	Sequence 7, Appl
32	34	53.1	137	3	US-09-456-830-15	Sequence 15, Appl
33	34	53.1	137	3	US-09-456-830-19	Sequence 19, Appl
34	34	53.1	137	3	US-09-456-830-37	Sequence 37, Appl
35	34	53.1	137	3	US-09-456-830-19	Sequence 37, Appl
36	34	53.1	137	4	US-09-002-285-7	Sequence 15, Appl
37	34	53.1	137	4	US-09-002-285-15	Sequence 19, Appl
38	34	53.1	137	4	US-09-002-285-37	Sequence 37, Appl
39	34	53.1	137	4	US-09-589-477-7	Sequence 7, Appl
40	34	53.1	137	4	US-09-589-477-15	Sequence 15, Appl
41	34	53.1	137	4	US-09-589-477-19	Sequence 19, Appl
42	34	53.1	137	4	US-09-589-477-37	Sequence 37, Appl
43	34	53.1	322	4	US-09-326-806-3	Sequence 3, Appl
44	34	53.1	490	1	US-08-361-611-2	Sequence 2, Appl
45	34	53.1	490	1	US-08-565-655-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13 Application US/08856841
Sequence 13, Appl
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MORTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
JOURNAL: GUTATHIONE S. TRANSFERASE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
US-08-856-841-13
Query Match 100.0%; Score 64; DB 3; Length 395;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTOOYTDQIE 12
Db 256 NKGTOOYTDQIE 267
RESULT 2
US-08-856-841-21
Sequence 21, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214, 658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 398
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 21:
US-08-856-841-21

Query Match 100.0%; Score 64; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
|||||
Db 245 NKGTOOYTDQIE 256

RESULT 3

US-08-856-841-20
; Sequence 20, Application US/088556841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 64; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
|||||
Db 256 NKGTOOYTDQIE 267

RESULT 4

US-08-856-841-22
; Sequence 22, Application US/088556841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543
; TYPE: AMINO ACID

```
Query Match          100.0%; Score 64; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches    12; Conservative   0; Indels    0; Gaps    0.
```

```

Db          597 NKGTQYTDQIE 608

RESULT 6
US-09-095-855-203
; Sequence 203, Application US/09095855
; Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTESD for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
Prior Application DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 203:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-855-203

Oy          1 NKGTQYTDQIE 12
Db          161 DKSTHTTNDID 172

Query Match          62.5%; Score 40; DB 3; Length 187;
Best Local Similarity 58.3%; Pred. No. 4.9;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 7
US-09-205-426-203
; Sequence 203, Application US/09205426
; Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Compounds and Methods for Treatment and
TITLE OF INVENTION: Diagnosis of Mycobacterial Infections

```

FILE REFERENCE: 11000.1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
9-205-426-203

Query Match 62.5%; Score 40; DB 4; Length 187;
Best Local Similarity 58.3%; Pred. No. 4.9;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOYDQIE 12
|||:|||||
Db 161 DKSTHQTND 172

RESULT 8
US-07-731-157A-2
Sequence 2, Application US/07731157A
Patent No. 5457032
GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misec, Onno
APPLICANT: Van der laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157A
FILING DATE: 19910509
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER PH.D.; BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GBRO-027/000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: AMINO ACID

TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-731-157A-2

Query Match 59.4%; Score 38; DB 1; Length 720;
Best Local Similarity 70.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTOOYDQIE 12
|||:|||||
Db 680 GTTHYSDQIE 689

RESULT 9
US-08-541-780-2
Sequence 2, Application US/08541780
Patent No. 5935831
GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misec, Onno
APPLICANT: Van der laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOLEY GODDARD CASTRO HUDDLESON & TATUM
STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,780
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/731,157
FILING DATE:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER PH.D.; BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GBRO-027/000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-541-780-2

Query Match 59.4%; Score 38; DB 2; Length 720;
Best Local Similarity 70.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTOOYDQIE 12
|||:|||||
Db 680 GTTHYSDQIE 689

RESULT 10
US-09-039-198A-14
Sequence 14, Application US/09039198A

Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-14

Query Match 57.8% Score 37; DB 3; Length 373;
Best Local Similarity 54.5% Pred. No. 37;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQI 11
DB 79 NFGTQKFTDMV 89

ULT 11
US-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 57.8% Score 37; DB 3; Length 373;
Best Local Similarity 54.5% Pred. No. 37;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQI 11
DB 79 NFGTQKFTDMV 89

RESULT 12
US-08-877-599-14
Sequence 14, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-877-599-14

Query Match 57.8% Score 37; DB 4; Length 373;
Best Local Similarity 54.5% Pred. No. 37;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQI 11

Db 79 NEGTOKFTDMV 89

RESULT 13
US-08-877

```

US-08-877-599-15
/ Sequence 15, Application US/08877599
/ Patent No. 6372212
/
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Patrick W.
/ TITLE OF INVENTION: Cellulase Materials and Methods
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 6300 Sears Tower, 233 South Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States of America
/ ZIP: 60606-6402
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/877,599
/ FILING DATE:
/
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/663,618
/ FILING DATE: 14-JUN-1996
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 27866/33994
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/
/ TELEX: 25-3856
/
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 373 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: peptide
/
/ US-08-877-599-15

```

	Very Match	Score 37;	DB 4;	Length 373;
	Local Similarity	54.5%	Pred. No. 37;	
	Matches	6;	Conservative	3; Mismatches 2; Indels 0; Gaps 0
QY	1 NKGTQOQYTDQI 11			
	--- --- ---			
	--- --- ---			
Db	79 NFGTQKPTDMV 89			

```

RESULT 14
US-09-267-574-14
; Sequence 14, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14

```

```

; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-267-574-14

```

Query Match	57.8%	Score 37	DB 4	Length 373	..
Best Local Similarity	54.5%	Pred. No. 37			
Matches	6	Conservative	3	Mismatches	2
				Indels	0
				Gaps	0

```

RESULT 15
US-09-267-574-15
: Sequence 15, Application US/09267574
: Patent No. 639571
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick M.
: TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
: FILE REFERENCE: 27866/35407
: CURRENT APPLICATION NUMBER: US/09/267,574
: CURRENT FILING DATE: 1999-03-12
: EARLIER APPLICATION NUMBER: 09/039,198
: EARLIER FILING DATE: 1998-03-12
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 15
: LENGTH: 373
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-267-574-15

```

Query Match	57.8%;	Score 37;	DB 4;	Length 373;
Best Local Similarity	54.5%;	Pred. No. 37;		
Matches	6;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0;

RESULT 16
 US-08-486-839-6
 Sequence 6, Application US/08486839
 Patent No. 5928928
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: A human chitinase, its recombinant
 production, its use for decomposing chitin, its use
 in therapy or prophylaxis against infection diseases
 TITLE OF INVENTION: In therapy or prophylaxis against infection diseases
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hoffmann & Baron
 STREET: 350 Jericho Turnpike
 CITY: Jericho
 STATE: New York
 COUNTRY: United States of America
 ZIP: 11758
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 SOFTWARE:
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/486, 839
 FILING DATE: 07 - June - 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Baron, Ronald J.
 REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-486-839-6

Query Match 57.8%; Score 37; DB 2; Length 387;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOQYTDQI 11
| | | | |
100 NFGTQKFTDMV 110

RESULT 17
US-09-151-011-6
Sequence 6, Application US/09151011
Patent No. 6057142
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A Human Chitinase, Its Recombinant
TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998

ATTORNEY/AGENT INFORMATION:

NAME: Morris, Robert C.

REGISTRATION NUMBER: 42,910

REFERENCE/DOCKET NUMBER: 294-32 DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-151-011-6

Query Match 57.8%; Score 37; DB 3; Length 387;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOQYTDQI 11
| | | | |
100 NFGTQKFTDMV 110

RESULT 18
US-09-343-623-6
Sequence 6, Application US/09343623
Patent No. 6303118
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
TITLE OF INVENTION: production, its use for decomposing chitin, its use
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-343-623-6

Query Match 57.8%; Score 37; DB 4; Length 387;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 NKGTOQYTDQI 11
| | | | |
100 NFGTQKFTDMV 110

RESULT 19
US-09-554-999-2
Sequence 2, Application US/09554999
Patent No. 6465186
GENERAL INFORMATION:

APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases from Gram Positive Organisms
FILE REFERENCE: GC389-US
CURRENT APPLICATION NUMBER: US/09/554,999
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: PCT/US98/27040
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: GB 9727471.6
PRIOR FILING DATE: 1997-12-30
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0

US-09-554-999-2

SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-554-999-2

Query Match 57.8%; Score 37; DB 4; Length 415;
Best Local Similarity 58.3%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQI 12
DB 208 NKGKPYTDQPE 219

RESULT 20
US-08-486-839-4
Sequence 4, Application US/08486839
Patent No. 5928928

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.

TITLE OF INVENTION: 16

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07 - June - 1995

ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: Protein
HYPOTHETICAL: NO

US-08-486-839-4

Query Match 57.8%; Score 37; DB 2; Length 466;
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQI 11
DB 100 NKGKPYTDQPE 110

RESULT 21
US-09-151-011-4
Sequence 4, Application US/09151011
Patent No. 6057142
GENERAL INFORMATION:
APPLICANT:

TITLE OF INVENTION: A Human Chitinase, Its Recombinant
production, its use for decomposing chitin, its use in
therapy or prophylaxis against infection diseases.

TITLE OF INVENTION: 16

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998

ATTORNEY/AGENT INFORMATION:
NAME: Morris, Robert C.
REGISTRATION NUMBER: 42,910
REFERENCE/DOCKET NUMBER: 294-32 DIV

TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: Protein
HYPOTHETICAL: NO

US-09-151-011-4

Query Match 57.8%; Score 37; DB 3; Length 466;
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQI 11
DB 100 NKGKPYTDQPE 110

RESULT 22
US-09-039-198A-2
Sequence 2, Application US/09039198A
Patent No. 6200951

GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 27866/34391
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 466 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-039-198A-2

Query Match 57.8%; Score 37; DB 4; Length 466;
 Best Local Similarity 54.5%; Pred. No. 47;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQI 11
 100 NFGTQKFTDMV 110

RESULT 23
 US-09-343-623-4
 Sequence 4, Application US/09343623
 Patent No. 6303118
 GENERAL INFORMATION:

APPLICANT:
 TITLE OF INVENTION: A human chitinase, its recombinant
 TITLE OF INVENTION: production, its use for decomposing chitin, its use
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hoffmann & Baron
 STREET: 350 Jericho Turnpike
 CITY: Jericho
 STATE: New York
 COUNTRY: United States of America
 ZIP: 11758

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 SOFTWARE:
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/343,623

FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/486,839
 FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:
 NAME: Baron, Ronald J.
 REGISTRATION NUMBER: 29,281
 REFERENCE/DOCKET NUMBER: 294-26
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (516) 822-3550
 TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 466 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 US-09-343-623-4

Query Match 57.8%; Score 37; DB 4; Length 466;
 Best Local Similarity 54.5%; Pred. No. 47;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQI 11

Db 100 NFGTQKFTDMV 110

RESULT 24
 US-08-877-599-2

Sequence 2, Application US/08877599
 Patent No. 6372212
 GENERAL INFORMATION:
 APPLICANT: Gray, Patrick W.
 TITLE OF INVENTION: Chitinase Materials and Methods
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/877,599
 FILING DATE:
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/663,618
 FILING DATE: 14-JUN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Rin-Laures, Li-Hsien
 REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 27866/33994
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 466 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-877-599-2

Query Match 57.8%; Score 37; DB 4; Length 466;
 Best Local Similarity 54.5%; Pred. No. 47;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQI 11
 Db 100 NFGTQKFTDMV 110

RESULT 25
 US-09-267-574-2

Sequence 2, Application US/09267574
 Patent No. 6399571
 GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
 APPLICANT: Tjoelker, Larry W.
 TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
 FILE REFERENCE: 27866/35407
 CURRENT APPLICATION NUMBER: US/09/267,574
 CURRENT FILING DATE: 1999-03-12
 EARLIER APPLICATION NUMBER: 09/039,198
 EARLIER FILING DATE: 1998-03-12
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 2
 LENGTH: 466

TYPE: PRT
ORGANISM: Homo sapiens
US-09-267-574-2

Query Match
Best Local Similarity 57.8%; Score 37; DB 4; Length 466;
Best Local Similarity 54.5%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 11
| : : : : :
DB 100 NKGTKYVDMV 110

RESULT 26
US-09-134-001C-4285
; Sequence 4285, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4285
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4285

Query Match
Best Local Similarity 56.2%; Score 36; DB 4; Length 738;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 12
| : : : : :
DB 213 NKGTKYVDMV 224

RESULT 27
US-09-238-303-11
; Sequence 11, Application US/09238303B
; Patent No. 6284253
; GENERAL INFORMATION:
; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6284253el Feline Immunodeficiency Virus Nucleotide Sequence
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/238,303B
; CURRENT FILING DATE: 1999-01-28
; EARLIER APPLICATION NUMBER: US 60/072,927
; EARLIER FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 11
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the env gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
; OTHER INFORMATION: Immunodeficiency virus
US-09-238-303-11

Query Match
Best Local Similarity 56.2%; Score 36; DB 3; Length 863;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GGTQYTDQIE 12
| : : : : :
DB 3 GGTQYTDQIE 12

DB 252 GLEQYVDQIQ 261

RESULT 28
US-09-946-239-11
; Sequence 11, Application US/09946239
; Patent No. 6579527
; GENERAL INFORMATION:
; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6579527el Feline Immunodeficiency Virus Nucleotide and
; TITLE OF INVENTION: Polypeptide Sequences
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/946,239
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 09/238,303, US 60/072,927
; PRIOR FILING DATE: 1999-01-28, 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 11
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the env gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
; OTHER INFORMATION: Immunodeficiency virus
US-09-946-239-11

Query Match
Best Local Similarity 56.2%; Score 36; DB 4; Length 863;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GGTQYTDQIE 12
| : : : : :
DB 252 GLEQYVDQIQ 261

RESULT 29
US-09-328-352-7978
; Sequence 7978, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7978
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7978

Query Match
Best Local Similarity 54.7%; Score 35; DB 4; Length 61;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 10
| : : : : :
DB 5 NKGTKYVDMV 14

RESULT 30
US-09-198-452A-753
; Sequence 753, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999

;; CURRENT APPLICATION NUMBER: US/09/198,452A
;; CURRENT FILING DATE: 1998-11-24
;; NUMBER OF SEQ ID NOS: 6849
;; SEQ ID NO 753
;; LENGTH: 91
;; TYPE: PRT
;; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-753

Query Match 54.7%; Score 35; DB 4; Length 91;
Best Local Similarity 54.5%; Pred. No. 19;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KGTQOYTDQIE 12
| | | | | : | |
| | | | | : | |
Db 31 KGVQOYLTLE 41

RESULT 31

US-09-456-830-7
Sequence 7, Application US/09456830
Patent No. 6274721

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest

APPLICANT: Wicker, Carol

APPLICANT: Narva, Kenneth E.

APPLICANT: Walz, Michelle

APPLICANT: Stockhoff

TITLE OF INVENTION: Toxins Active Against Pests

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Saliwanchik & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/456,830

FILING DATE: 07-DEC-1999

CLASSIFICATION: 534

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: MA701

TELECOMMUNICATION INFORMATION:

TELEPHONE: (352) 375-8100

TELEFAX: (352) 372-5800

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 137 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-456-830-7

Qy 1 NKGTQOYTDQIE 12
| | | | | : | |
| | | | | : | |
Db 101 NSGNEVYIDRIE 112

Query Match 53.1%; Score 34; DB 3; Length 137;
Best Local Similarity 50.0%; Pred. No. 45;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQIE 12
| | | | | : | |
| | | | | : | |
Db 101 NSGNEVYIDRIE 112

RESULT 32
US-09-456-830-15

;; Sequence 15, Application US/09456830

;; Patent No. 6274721

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest

APPLICANT: Wicker, Carol

APPLICANT: Narva, Kenneth E.

APPLICANT: Walz, Michelle

APPLICANT: Stockhoff

TITLE OF INVENTION: Toxins Active Against Pests

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Saliwanchik & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/456,830

FILING DATE: 07-DEC-1999

CLASSIFICATION: 534

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: MA701

TELECOMMUNICATION INFORMATION:

TELEPHONE: (352) 375-8100

TELEFAX: (352) 372-5800

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 137 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-456-830-15

Qy 1 NKGTQOYTDQIE 12
| | | | | : | |
| | | | | : | |
Db 101 NSGNEVYIDRIE 112

RESULT 33

US-09-456-830-19

Sequence 19, Application US/09456830

Patent No. 6274721

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest

APPLICANT: Wicker, Carol

APPLICANT: Narva, Kenneth E.

APPLICANT: Walz, Michelle

APPLICANT: Stockhoff

TITLE OF INVENTION: Toxins Active Against Pests

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Saliwanchik & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

MOLECULE TYPE: protein;

RESOL 36
US-09-002-285-15

Sequence 15, Application US/09002285
Patent No. 6369213
GENERAL INFORMATION:
APPLICANT: Schneck, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002,285
FILING DATE:
CLASSIFICATION:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
9-002-285-15

Query Match 53.1%; Score 34; DB 4; Length 137;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
DB 101 NSGNEVYIDRIE 112

RESULT 37
US-09-002-285-19
Sequence 19, Application US/09002285
Patent No. 6369213
GENERAL INFORMATION:
APPLICANT: Schneck, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002,285
FILING DATE:
CLASSIFICATION:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-002-285-19

Query Match 53.1%; Score 34; DB 4; Length 137;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
DB 101 NSGNEVYIDRIE 112

RESULT 38
US-09-002-285-37
Sequence 37, Application US/09002285
Patent No. 6369213
GENERAL INFORMATION:
APPLICANT: Schneck, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002,285
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
9-002-285-37

Query Match 53.1%; Score 34; DB 4; Length 137;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTOGYTDQIE 12
DB 101 NSGNEVYIDRIE 112

RESULT 39

US-09-589-477-7
Sequence 7, Application US/09589477
Patent No. 6570005
GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/589,477
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C1C1
TELECOMMUNICATION INFORMATION:

TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-589-477-7

Query Match 53.1%; Score 34; DB 4; Length 137;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTOGYTDQIE 12
DB 101 NSGNEVYIDRIE 112

RESULT 40

US-09-589-477-15
Sequence 15, Application US/09589477
Patent No. 6570005
GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/589,477
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-589-477-15
Query Match 53.1%; Score 34; DB 4; Length 137;

Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches

4; Indels 0; Gaps 0;

Qy 1 NKGTOOYTDQIE 12
Db 101 NSGNEVYIDRIE 112

Search completed: August 20, 2003, 09:34:27
Job time: 13.2371 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:30:58 ; Search time 48.7423 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-4
Perfect score: 64
Sequence: 1 NKG2QYTDQIE 12

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 497079 seqs, 131961718 residues
1 number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCITUS_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCITUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	756	US-10-205-942-4	Sequence 4, Appli
2	40	62.5	187	US-10-205-979-49	Sequence 49, Appl
3	40	62.5	187	US-10-051-643-203	Sequence 203, Appl
4	37	57.8	233	US-10-188-586A-7	Sequence 7, Appli
5	37	57.8	373	US-10-161-547-14	Sequence 14, Appl
6	37	57.8	373	US-10-161-547-15	Sequence 15, Appl
7	37	57.8	415	US-10-188-586A-2	Sequence 2, Appli
8	37	57.8	445	US-10-004-219B-10	Sequence 10, Appl
9	37	57.8	466	US-10-161-547-2	Sequence 2, Appli
10	36	56.2	92	US-10-106-698-6596	Sequence 6596, Ap
11	36	56.2	863	US-09-946-239-11	Sequence 11, Appl
12	35	54.7	428	US-09-347-331-10	Sequence 10, Appl
13	35	54.7	428	US-09-563-728A-28	Sequence 28, Appl
14	34	53.1	137	US-10-099-285-7	Sequence 7, Appli
15	34	53.1	137	US-10-099-285-15	Sequence 15, Appl

16	34	53.1	137	US-10-099-285-19	Sequence 19, Appl
17	34	53.1	137	US-10-099-285-37	Sequence 37, Appl
18	34	53.1	137	US-09-870-759-82	Sequence 82, Appl
19	34	53.1	438	US-10-043-487-175	Sequence 175, Appl
20	34	53.1	549	US-10-097-065-139	Sequence 139, Appl
21	34	53.1	557	US-10-199-672-326	Sequence 326, Appl
22	34	53.1	557	US-10-187-749-326	Sequence 326, Appl
23	34	53.1	557	US-10-194-457-326	Sequence 326, Appl
24	34	53.1	557	US-10-052-586-326	Sequence 326, Appl
25	34	53.1	557	US-10-174-590-326	Sequence 326, Appl
26	34	53.1	557	US-10-176-758-326	Sequence 326, Appl
27	34	53.1	557	US-10-175-737-326	Sequence 326, Appl
28	34	53.1	557	US-10-173-706-326	Sequence 326, Appl
29	34	53.1	557	US-10-175-738-326	Sequence 326, Appl
30	34	53.1	557	US-10-175-752-326	Sequence 326, Appl
31	34	53.1	557	US-10-176-482-326	Sequence 326, Appl
32	34	53.1	557	US-10-176-757-326	Sequence 326, Appl
33	34	53.1	557	US-10-176-913-326	Sequence 326, Appl
34	34	53.1	557	US-10-180-552-326	Sequence 326, Appl
35	34	53.1	557	US-10-180-557-326	Sequence 326, Appl
36	34	53.1	557	US-10-173-700-326	Sequence 326, Appl
37	34	53.1	557	US-10-174-572-326	Sequence 326, Appl
38	34	53.1	557	US-10-174-579-326	Sequence 326, Appl
39	34	53.1	557	US-10-174-582-326	Sequence 326, Appl
40	34	53.1	557	US-10-174-588-326	Sequence 326, Appl
41	34	53.1	557	US-10-175-739-326	Sequence 326, Appl
42	34	53.1	557	US-10-175-740-326	Sequence 326, Appl
43	34	53.1	557	US-10-175-743-326	Sequence 326, Appl
44	34	53.1	557	US-10-176-488-326	Sequence 326, Appl
45	34	53.1	557	US-10-176-492-326	Sequence 326, Appl

ALIGNMENTS

RESULT 1
US-10-205-942-4
; Sequence 4, Application US/10205942
; Publication No. US20030053990A1
GENERAL INFORMATION:
; APPLICANT: University of No. US20030053990A1Ch Carolina-Chapel Hill
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/10/205,942
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Adeno-associated virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) ..(2271)
; OTHER INFORMATION: B19/AAV chimeric capsid coding sequence
US-10-205-942-4
Query Match 100.0%; Score 64; DB 15; Length 756;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKG2QYTDQIE 12
DB 597 NKG2QYTDQIE 608
RESULT 2
US-10-205-979-49
; Sequence 49, Application US/10205979
; Publication No. US20030147861A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Tan, Paul L. J.
/ APPLICANT: Abernethy, Nevin
/ TITLE OF INVENTION: Compounds and Methods for the Modulation
/ FILE REFERENCE: 11000.1063U
/ CURRENT FILING DATE: 2002-07-25
/ PRIOR FILING DATE: 2001-07-26
/ PRIOR APPLICATION NUMBER: 60/308,446
/ NUMBER OF SEQ ID NOS: 52
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 49
/ LENGTH: 187
/ TYPE: PRT
/ ORGANISM: Mycobacterium vaccae
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(187)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-205-979-49
```

```
Query Match      62.5%; Score 40; DB 12; Length 187;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 NKGTQOYTDQIE 12
       :|:|:|:|:|:|:
Db      161 DKSTHQYTNQID 172
```

```
RESULT 3
/ Sequence 203, Application US/10051643
/ Publication No. US20020197265A1
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Tan, Paul L. J.
/ TITLE OF INVENTION: Methods and Compounds for the Treatment
/ TITLE OF INVENTION: System using Mycobacterium vaccae
/ FILE REFERENCE: 11000.1008C2
/ CURRENT APPLICATION NUMBER: US/10/051,643
/ CURRENT FILING DATE: 2002-01-18
/ PRIOR FILING DATE: 1998-09-17
/ PRIOR APPLICATION NUMBER: US09/156,181
/ PRIOR FILING DATE: 1997-12-23
/ PRIOR APPLICATION NUMBER: US 08/996,624
/ NUMBER OF SEQ ID NOS: 208
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 203
/ LENGTH: 187
/ TYPE: PRT
/ ORGANISM: Mycobacterium vaccae
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: (186)
US-10-051-643-203
```

```
Query Match      62.5%; Score 40; DB 14; Length 187;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 NKGTQOYTDQIE 12
       :|:|:|:|:|:|:
Db      161 DKSTHQYTNQID 172
```

```
RESULT 4
/ Sequence 7, Application US/10188586A
/ Publication No. US20030045438A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Batell, David A.
/ TITLE OF INVENTION: Proteases From Gram Positive Organisms
/ FILE REFERENCE: GC389-US-D1
/ CURRENT APPLICATION NUMBER: US/10/188,586A
/ CURRENT FILING DATE: 2002-07-02
/ PRIOR APPLICATION NUMBER: US 09/554,999
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: PCT/US98/27040
/ PRIOR FILING DATE: 1998-12-17
/ PRIOR APPLICATION NUMBER: GB 9727471.6
/ PRIOR FILING DATE: 1997-12-30
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 233
/ TYPE: PRT
/ ORGANISM: Bacillus subtilis
US-10-188-586A-7
```

```
Query Match      57.8%; Score 37; DB 15; Length 233;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 NKGTQOYTDQIE 12
       :|:|:|:|:|:|:
Db      206 NQKKKPYTDQPE 217
```

```
RESULT 5
/ Sequence 14, Application US/10161547
/ Publication No. US20030143216A1
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Patrick W.
/ APPLICANT: Tjoelker, Larry W.
/ TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
/ FILE REFERENCE: 27866/35407
/ CURRENT APPLICATION NUMBER: US/10/161,547
/ CURRENT FILING DATE: 2002-06-03
/ PRIOR APPLICATION NUMBER: US/09/267,574
/ PRIOR FILING DATE: 1999-03-12
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 373
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-161-547-14
```

```
Query Match      57.8%; Score 37; DB 12; Length 373;
Best Local Similarity 54.5%; Pred. No. 79;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 NKGTQOYTDQIE 11
       :|:|:|:|:|:|:
Db      79 NFGTQKPYTDQV 89
```

```
RESULT 6
/ Sequence 15, Application US/10161547
/ Publication No. US20030143216A1
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Patrick W.
/ APPLICANT: Tjoelker, Larry W.
/ TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
/ FILE REFERENCE: 27866/35407
/ CURRENT APPLICATION NUMBER: US/10/161,547
/ CURRENT FILING DATE: 2002-06-03
/ PRIOR APPLICATION NUMBER: US/09/267,574
```


PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-547-15

Query Match 57.8%; Score 37; DB 12; Length 373;
Best Local Similarity 54.5%; Pred. No. 79;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQI 11
DB 79 NFGTQKFTDMV 89

LT 7
0-188-586A-2

Sequence 2, Application US/10188586A
Publication No. US20030045438A1
GENERAL INFORMATION:
APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases From Gram Positive Organisms
FILE REFERENCE: GC389-US-D1
CURRENT APPLICATION NUMBER: US/10/188,586A
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 09/554,999
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: PCT/US98/27040
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: GB 9727471.6
PRIOR FILING DATE: 1997-12-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-188-586A-2

Query Match 57.8%; Score 37; DB 15; Length 415;
Best Local Similarity 58.3%; Pred. No. 89;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQI 12
DB 208 NQKKFYTDQPE 219

RESULT 8
US-10-004-219B-10
Sequence 10, Application US/10004219B
Publication No. US20030087414A1
GENERAL INFORMATION:

APPLICANT: Macrozyme
APPLICANT: Aerts, Johannes M.F.G.
TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
FILE REFERENCE: 2183-5136US
CURRENT APPLICATION NUMBER: US/10/004,219B
CURRENT FILING DATE: 2001-11-02
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SITE
LOCATION: (1) ..(445)
OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

Query Match 57.8%; Score 37; DB 15; Length 445;
Best Local Similarity 54.5%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQI 11
DB 79 NFGTQKFTDMV 89

RESULT 9
US-10-161-547-2
Sequence 2, Application US/10161547
Publication No. US20030143216A1
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/10/161,547
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/267,574
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-547-2

Query Match 57.8%; Score 37; DB 12; Length 466;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOYTDQI 11
DB 100 NFGTQKFTDMV 110

RESULT 10
US-10-106-6596-6596
Sequence 6596, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 6596
LENGTH: 92
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (76)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-106-698-6596

Query Match 56.2%; Score 36; DB 15; Length 92;

Best Local Similarity 60.0%; Pred. No. 26;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTGQYTDQIE 12
|:|:|:|:|:
DB 30 GTGQYTDQIE 39

RESULT 11

US-09-946-239-11

Sequence 11, Application US/09946239

Patent No. US20020044945A1

GENERAL INFORMATION:

APPLICANT: Barr, Margaret C.

TITLE OF INVENTION: No. US20020044945A1 Feline Immunodeficiency Virus Nucleotide an

FILE REFERENCE: 18617.0059

CURRENT FILING DATE: 2001-09-04

PRIOR FILING DATE: 1999-01-28, 1998-01-29

NUMBER OF SEQ ID NOS: 17

SEQ ID NO 11

LENGTH: 863

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: protein encoded by the env gene of a recombinant viral

OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas' cat feline

US-09-946-239-11

Query Match 56.2%; Score 36; DB 9; Length 863;

Best Local Similarity 60.0%; Pred. No. 3e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTGQYTDQIE 12
|:|:|:|:|:
DB 252 GLEQYVDQIQ 261

RESULT 12

US-09-347-331-10

Sequence 10, Application US/09347331

Patent No. US20010010909A1

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.

APPLICANT: Miao, Gou Hau

APPLICANT: Odell, Joan

APPLICANT: Sakai, Hajime

TITLE OF INVENTION: Chromatin Associated Proteins

FILE REFERENCE: BB-1118-C

CURRENT APPLICATION NUMBER: US/09/347,331

CURRENT FILING DATE: 1999-07-02

EARLIER APPLICATION NUMBER: 06/092,841

EARLIER FILING DATE: July 14, 1998

NUMBER OF SEQ ID NOS: 11

SOFTWARE: Microsoft Office 97

SEQ ID NO 10

LENGTH: 428

TYPE: PRT

ORGANISM: Gallus gallus

US-09-347-331-10

Query Match 54.7%; Score 35; DB 9; Length 428;

Best Local Similarity 54.5%; Pred. No. 2.1e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 11
|:|:|:|:|:
DB 1 NKGTQYTDQI 11

DB 348 NQNSROYLDQI 358

RESULT 13

US-09-563-728A-28

Sequence 28, Application US/09563728A

Publication No. US20030078216A1

GENERAL INFORMATION:

APPLICANT: MacLeod, Alan R

APPLICANT: Li, Zoumet

APPLICANT: Besteman, Jeffrey M

TITLE OF INVENTION: Inhibition of Histone Deacetylase

FILE REFERENCE: 106101.229

CURRENT APPLICATION NUMBER: US/09/563,728A

CURRENT FILING DATE: 2000-05-03

PRIOR APPLICATION NUMBER: 60/132,287

PRIOR FILING DATE: 1999-05-03

NUMBER OF SEQ ID NOS: 36

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 28

LENGTH: 428

TYPE: PRT

ORGANISM: Homo sapiens

US-09-563-728A-28

Query Match 54.7%; Score 35; DB 11; Length 428;

Best Local Similarity 54.5%; Pred. No. 2.1e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 11
|:|:|:|:|:
DB 348 NQNSROYLDQI 358

RESULT 14

US-10-099-285-7

Sequence 7, Application US/10099285

Publication No. US20030105319A1

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest

APPLICANT: Wicker, Carol E.

APPLICANT: Narva, Kenneth E.

APPLICANT: Walz, Michelle

APPLICANT: Stockhoff, Brian

APPLICANT: Muller-Cohn, Judy

TITLE OF INVENTION: Toxins Active Against Pests

NUMBER OF SEQUENCES: 105

CORRESPONDENCE ADDRESS:

ADDRESS: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/099,285

FILING DATE: 15-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/002,285

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/886,615

FILING DATE: 1-JUL-1997

APPLICATION NUMBER: US 08/674,002

FILING DATE: 1-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

RESULT 17
US-10-099-285-37
Sequence 37, Application US/10099285
Publication No. US20030105319A1
GENERAL INFORMATION:
APPLICANT: Schneck, H. Ernest
Mickler, Carol
Narva, Kenneth E.
Walz, Michelle
Stockhoff, Brian
Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/099,285
FILING DATE: 15-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/002,285
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-10-099-285-37
Query Match 53.1%; Score 34; DB 15; Length 137;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQIE 12
DB 101 NKGNEVYIDRIE 112
RESULT 18
US-09-870-759-82
Sequence 82, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: Termán, David S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870,759
CURRENT FILING DATE: 2002-01-14

PRIOR APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: Patentin version 3.1
SEQ ID NO 82
LENGTH: 330
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-870-759-82
Query Match 53.1%; Score 34; DB 10; Length 330;
Best Local Similarity 41.7%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQIE 12
DB 231 NKGTFEYDMWK 242
RESULT 19
US-10-043-487-375
Sequence 375, Application US/10043487
Publication No. US20030055220A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: Pierre, LÉGRAIN
TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptid
FILE REFERENCE: B4778A
CURRENT APPLICATION NUMBER: US/10/043,487
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/261,130
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 561
SOFTWARE: Patentin version 3.1
SEQ ID NO 375
LENGTH: 438
TYPE: PRT
ORGANISM: Shigella Flexneri
US-10-043-487-375
Query Match 53.1%; Score 34; DB 15; Length 438;
Best Local Similarity 54.5%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 KGTQOYTDQIE 12
DB 128 KVTGEYPPDVE 138
RESULT 20
US-10-097-065-139
Sequence 139, Application US/10097065
Publication No. US20030055236A1
GENERAL INFORMATION:
APPLICANT: Moore, Paul A. et al.
TITLE OF INVENTION: 110 Human Secreted Proteins
FILE REFERENCE: P2021PI.
CURRENT APPLICATION NUMBER: US/10/097,065
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: PCT/US98/27059
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,007
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,057
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,006
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,369
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,367

PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,368
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,169
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,008
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,365
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 139
LENGTH: 549
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (132)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (398)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-097-065-139

Query Match 53.1%; Score 34; DB 15; Length 549;
Best Local Similarity 50.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
DB 406 NKATQLFNTQTD 417

RESULT 21
US-10-199-672-326
Sequence 326, Application US/10199672
Publication No. US20030148442A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/199,672
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: US/10/052,586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-199-672-326

Query Match 53.1%; Score 34; DB 12; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
DB 414 NKATQLFNTQTD 425

RESULT 22
US-10-187-749-326
Sequence 326, Application US/10187749
Publication No. US20030153036A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/187,749
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: US/10/052,586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-187-749-326

PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/08623
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908

Query Match 53.1% Score 34; DB 14; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDJLE 12
DB 414 NKATQJFTNQTD 425

RESULT 25
US-10-174-590-326
Sequence 326, Application US/10174590
Publication No. US20030008352A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Deanovers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C42
CURRENT APPLICATION NUMBER: US/10/174,590
CURRENT FILING DATE: 2002-06-18
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-174-590-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQFTNQTD 425

RESULT 26

US-10-176-758-326
Sequence 326, Application US/10176758
Publication No. US2003008353A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/176,758
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-758-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQFTNQTD 425

RESULT 27

US-10-175-737-326
Sequence 326, Application US/10175737
Publication No. US2003003153A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/175,737
CURRENT FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-326

APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C50
CURRENT APPLICATION NUMBER: US/10/175,737
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQFTNQTD 425

RESULT 28

US-10-173-706-326
Sequence 326, Application US/10173706
Publication No. US2003002293A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
CURRENT FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQFTNQTD 425

RESULT 29

US-10-175-738-326
Sequence 326, Application US/10175738
Publication No. US20030022294A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/175,738
CURRENT FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-738-326

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-738-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
DB 414 NKATQLFNTQTD 425

RESULT 30
US-10-175-752-326
Sequence 326, Application US/10175752
Publication No. US20030022295A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C60
CURRENT APPLICATION NUMBER: US/10/175,752
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-752-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
DB 414 NKATQLFNTQTD 425

RESULT 31
US-10-176-482-326
Sequence 326, Application US/10176482
Publication No. US20030022296A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C70
CURRENT APPLICATION NUMBER: US/10/176,482
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-482-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
DB 414 NKATQLFNTQTD 425

RESULT 32
US-10-176-757-326
Sequence 326, Application US/10176757
Publication No. US20030022297A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C86
CURRENT APPLICATION NUMBER: US/10/176,757
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 326
LENGTH: 557
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-757-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
DB 414 NKATQLFNTQTD 425

RESULT 33
US-10-176-482-326
Sequence 326, Application US/10176482
Publication No. US20030022296A1

US-10-176-913-326
; Sequence 326, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 414 NKATQLEFNTQTD 425

RESULT 34
US-10-180-552-326
; Sequence 326, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 414 NKATQLEFNTQTD 425

RESULT 35
US-10-180-557-326
; Sequence 326, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 414 NKATQLEFNTQTD 425

RESULT 36
US-10-173-700-326
; Sequence 326, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-700-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQLFNTQTD 425

RESULT 37

US-10-174-572-326
; Sequence 326, Application US/10174572
; Publication No. US20030027263A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C40
; CURRENT APPLICATION NUMBER: US/10/174,572
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-572-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQLFNTQTD 425

RESULT 38

US-10-174-579-326
; Sequence 326, Application US/10174579
; Publication No. US20030027264A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-579-326

Query Match 53.1%; Score 34; DB 15; Length 557;

Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQLFNTQTD 425

RESULT 39

US-10-174-582-326
; Sequence 326, Application US/10174582
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C36
; CURRENT APPLICATION NUMBER: US/10/174,582
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-582-326

Query Match 53.1%; Score 34; DB 15; Length 557;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
|||:|:|:
Db 414 NKATQLFNTQTD 425

RESULT 40

US-10-174-588-326
; Sequence 326, Application US/10174588
; Publication No. US20030027266A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C28
; CURRENT APPLICATION NUMBER: US/10/174,588
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 326
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-174-588-326

Query Match

53.1%; Score 34; DB 15; Length 557;

Best Local Similarity

50.0%; Pred. No. 4.3e+02;

Matches

6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQIE 12

Db 414 NKATQLEFTNQTD 425

Search completed: August 20, 2003, 10:08:17
Job time : 49.7423 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:07 ; Search time 13.4845 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-4
Perfect score: 64
Sequence: 1 NKGTOQYTDQIE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	781	1 JCVPI9	coat protein VP1 -
2	43	67.2	547	2 JCVPI9	coat protein VP1 -
3	43	67.2	548	2 JCVPI9	coat protein VP1 -
4	38	59.4	36	2 A60343	transforming prote
5	38	59.4	3971	2 T44806	mycosubtilin synth
6	37	57.8	169	2 D66801	prophage p13 prote
7	37	57.8	238	2 S11779	trypsin (BC 3.4.21
8	37	57.8	338	2 A53065	CCHAT enhancer-bin
9	37	57.8	415	2 G69885	processing protein
10	37	57.8	505	2 G46642	DNA primase (EC 2.
11	37	57.8	509	2 S45631	polyprotein chain
12	37	57.8	2475	2 S35307	polyprotein PP220
13	36	56.2	310	2 T02202	2'-hydroxyisoflavo
14	36	56.2	367	2 B84404	aminopeptidase hom
15	36	56.2	414	2 H85512	hypothetical prote
16	36	56.2	581	2 H69452	hypothetical prote
17	36	56.2	906	2 G75362	aconitase hydratase
18	35	56.2	1215	2 I52882	autocatalytic - huma
19	35	56.2	84	2 C86579	CT666 hypotetical
20	35	54.7	84	2 F72046	conserved hypotet
21	35	54.7	375	2 H90786	hypothetical prote
22	35	54.7	375	2 H85646	hypothetical prote
23	35	54.7	414	1 QDECA9	yaaf protein, 49k
24	35	54.7	414	2 B90662	hypothetical prote
25	35	54.7	424	2 JCVI02	histone deacetylase
26	35	54.7	428	2 JCVI02	histone deacetylase
27	35	54.7	516	2 F82070	2-Isopropylmalate
28	35	54.7	677	2 G87586	cation transportin
29	35	54.7	714	2 S76082	hypothetical prote

30	34.5	53.9	1270	2 T22615	hypothetical prote
31	34	53.1	106	2 A69872	conserved hypotet
32	34	53.1	120	2 D69668	hypothetical prote
33	34	53.1	246	2 H82553	3-demethylubiquino
34	34	53.1	255	2 H69128	conserved hypotet
35	34	53.1	270	2 T03635	plasticid-lipid-aso
36	34	53.1	289	2 B89991	truncated beta-hem
37	34	53.1	299	2 A63278	hypothetical prote
38	34	53.1	304	2 H75378	hypothetical prote
39	34	53.1	310	2 T04905	fibrillin homolog
40	34	53.1	318	2 T01472	probable fibrillin
41	34	53.1	322	2 S56633	plant fibrillin pr
42	34	53.1	327	2 D62090	conserved hypotet
43	34	53.1	331	2 S15324	beta-hemolysin - S
44	34	53.1	375	2 T03325	gene 118 protein -
45	34	53.1	466	2 D84906	probable beta-keto

ALIGNMENTS

```
RESULT 1
VCPV19
coat protein VP1 - parvovirus B19 (strain Au)
C:Species: parvovirus B19
A:Note: host Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: A24299
R:Shade, R.O., Blundell, M.C., Cotmore, S.F., Tattersall, P., Astell, C.R.
J. Virol. 58, 921-936, 1986
A>Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr
A:Reference number: A24299; MUID:86200451; PMID:3701931
A:Accession: A24299
A:Molecule type: DNA
A:Residues: 1-781 <SHA>
A:Cross-References: EMBL:ML1178; NID:G333375; PIDN:AAA66867.1; PID:G333377
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match      100.0%; Score 64; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 NKGTOQYTDQIE 12
Db      622 NKGTOQYTDQIE 633

RESULT 2
JCV4519
heat-shock protein groEL - Pasteurella multocida
C:Species: Pasteurella multocida
C>Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C:Accession: JCV4519
R:Love, B.C., Hansen, L.M., Hirsch, D.C.
Gene 166, 179-180, 1995
A>Title: Cloning and sequence of the groEL heat-shock operon of Pasteurella multocida.
A:Reference number: JCV4519; MUID:96105224; PMID:8525887
A:Accession: JCV4519
A:Molecule type: DNA
A:Residues: 1-547 <LOV>
A:Cross-References: GB:U0165; NID:G1144300; PIDN:AAA84916.1; PID:G1144302
A:Experimental source: serotype A:3
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; stress-induced protein

Query Match      67.2%; Score 43; DB 2; Length 547;
Best Local Similarity 66.7%; Pred. No. 3.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

Db 479 NAGTEQYGMIE 490

RESULT 3

C64076

Chaperonin groEL - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999

C:Accession: C64076

R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

J.; Gockyne, U.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

D.M.; Brandon, R.C.; Fine, L.D.; Fitchman, J.L.; Geophagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: C64076

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-548 <TRIG>

A:Cross-references: GB:U32736; GB:U42023; NID:91573519; PIDN:AAK22201.1; PID:91573528; T

C:Genetics:

A:Gene: groEL

C:Superfamily: chaperonin groEL

C:Keywords: heat shock; molecular chaperone; stress-induced protein

Query Match 67.2%; Score 43; DB 2; Length 548;

Best Local Similarity 66.7%; Pred. No. 3.9;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12

Db 480 NAGTEQYGMIE 491

RESULT 4

A60343

transforming protein (r-fos) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 07-May-1995

C:Accession: A60343

R:Cochran, B.H.; Zullo, J.; Verma, I.M.; Stiles, C.D.

Science 226, 1080-1082, 1984

A:Title: Expression of the c-fos gene and of an fos-related gene is stimulated by platelet

A:Reference number: A60343; MUID:85040409; PMID:6093261

A:Accession: A60343

A:Molecule type: mRNA

A:Residues: 1-36 <COC>

A:Cross-references: GB:K02785

A:Superfamily: fos transforming protein; fos/jun DNA-binding domain homology

C:Keywords: DNA binding; transforming protein

Query Match 59.4%; Score 38; DB 2; Length 36;

Best Local Similarity 63.6%; Pred. No. 1.7;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KGTQYTDQIE 12

Db 15 KGTETDQIE 25

RESULT 5

T44806

mycosubtilin synthetase chain myca [imported] - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Nov-2000

C:Accession: T44806

R:Dutman, E.H.; Hancock, L.W.; Rembold, M.; Venema, G.; Seltz, H.; Saenger, W.; Bernhard

Proc. Natl. Acad. Sci. U.S.A. 96, 13294-13299, 1999

A:Title: The mycosubtilin synthetase of Bacillus subtilis ATCC6633: A multifunctional hy

A:Reference number: Z22848; MUID:20027541; PMID:10557314

A:Accession: T44806

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3971 <DUI>

A:Cross-references: EMBL:AF184956; NID:96449053; PIDN:AAF08795.1; PID:96449055

A:Experimental source: strain ATCC6633

C:Genetics:

A:Gene: myca

C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase

C:Keywords: antibiotic biosynthesis; carrier protein

F:56-548/Domain: acetate-CoA ligase homology <ACLI>

F:581-648/Domain: acyl carrier protein homology <ACPI>

F:690-1050/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>

F:1290-1361/Domain: acyl carrier protein homology <ACP2>

F:2408-2476/Domain: acyl carrier protein homology <ACP3>

F:2986-3427/Domain: acetate-CoA ligase homology <ACI2>

F:3445-3513/Domain: acyl carrier protein homology <ACP4>

Query Match 59.4%; Score 38; DB 2; Length 3971;

Best Local Similarity 66.7%; Pred. No. 3.2e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQYTDQI 10

Db 1496 KSGKQYTDI 1504

RESULT 6

D86801

prophage p13 protein 39 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: D86801

R:Boletín, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrli

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: D86801

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-169 <STO>

A:Cross-references: GB:AE005176; PID:912724400; PIDN:AAK05510.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: p1339

Query Match 57.8%; Score 37; DB 2; Length 169;

Best Local Similarity 54.5%; Pred. No. 15;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 11

Db 60 NNSQGYADQL 70

RESULT 7

S31779

trypsin (EC 3.4.21.4) III precursor - Atlantic salmon (fragment)

C:Species: Salmo salar (Atlantic salmon)

C>Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999

C:Accession: S66657; S31779

R:Male, R.; Lorens, J.B.; Smalae, A.O.; Torrisen, K.R.

Eur. J. Biochem. 232, 677-685, 1995

A:Title: Molecular cloning and characterization of anionic and cationic variants of tryp

A:Reference number: S66657; MUID:9603508; PMID:7556223

A:Accession: S66657

A:Molecule type: mRNA

A:Residues: 1-238 <MAL>

A:Cross-references: EMBL:X70074; NID:964387; PIDN:CAA49679.1; PID:964388

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine protease

F:1-7/Domain: signal sequence (fragment) #status predicted <SIG>

F:8-15/Domain: activation peptide #status predicted <AP>

F:16-238/Product: trypsin III #status predicted <MAT>

F,16-231/Domain: trypsin homology <TR>
F,22-152,40-56,124-225,131-198,163-177,186-212/disulfide bonds: #status predicted
F,55,99,192/Active site: His, Asp, Ser #status predicted

Query Match 57.8% Score 37; DB 2; Length 238;
Best Local Similarity 41.7% Pred. No. 22;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQGYTDQIE 12
Db 73 NKGTEQFIDSVK 84

RESULT 8

CCAT enhancer-binding protein - California sea hare

N/Alternate names: C/BBP
C/Species: Aplysia californica (California sea hare)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000

A/Accession: A53066
Accession: A53066

A/Title: C/BBP is an immediate-early gene required for the consolidation of long-term fa
Cell 76, 1099-1114, 1994

A/Reference number: A53066; MUID:94185169; PMID:8137425

A/Accession: A53066
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-338 <ALB>
A/Cross-references: GB:U00994; NID:G392976; PIDN:AAA18286.1; PID:G487963

Query Match 57.8% Score 37; DB 2; Length 338;
Best Local Similarity 66.7% Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQGYTDQ 10
Db 258 KGTQGYVDK 266

RESULT 9

processing proteinase homolog ymfH - Bacillus subtilis

C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C/Accession: G69885

R/Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C/Bron, S.; Broutlet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Chd
Schlich, S.D.; Emerson, P.T.; Emtian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Cell 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Gallizi, A.; Gall
lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogilvay, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A/Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiuchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpeira, P.; Tognoni, A.; Tosco, V.; Uchiyama,
T.; Winers, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: G69885

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-415 <KUN>

A/Cross-references: GB:Z99112; GB:AL009126; NID:G2633902; PIDN:CAE13559.1; PID:G2634058

A/Experimental source: strain 168

C/Genetics:

A/Supersfamily: mitochondrial processing peptidase alpha chain

Query Match 57.8% Score 37; DB 2; Length 415;
Best Local Similarity 58.3% Pred. No. 40;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 NKGTQGYTDQIE 12
Db 208 NKGKPYTDQPE 219

RESULT 10

DNA primase (EC 2.7.7.-) 54K chain - mouse

N/Alternate names: DNA polymerase alpha/DNA primase complex 54K chain

C/Species: Mus musculus (house mouse)

C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000

C/Accession: C46642; S45629

R/Miyazawa, H.; Izumi, M.; Tada, S.; Takada, R.; Masutani, M.; Ui, M.; Hanaoka, F.
J. Biol. Chem. 268, 8111-8122, 1993

A/Title: Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase al
A/Reference number: A46642; MUID:93216788; PMID:8463324

A/Accession: C46642

A/Status: preliminary
A/Molecule type: protein

A/Residues: 1-505 <MTY>

A/Cross-references: GB:D13545; NID:G303662; PIDN:BA02745.1; PID:G303663

A/Experimental source: FM3A cells

A/Note: sequence extracted from NCBI backbone (NCBIN:129150, NCBI:129151)
R/Stadlbauer, F.; Bruckner, A.; Rehfuss, C.; Eckerskorn, C.; Lottspeich, F.; Foerster,
Eur. J. Biochem. 222, 781-793, 1994

A/Title: DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.
A/Reference number: S45628; MUID:94298818; PMID:8026492

A/Accession: S45629

A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-108, 'G', 110-505 <STA>

A/Cross-references: EMBL:D13385; NID:G397830; PIDN:BA04203.1; PID:G442472

C/Supersfamily: mouse DNA primase 54K chain

C/Keywords: DNA binding; nucleotidyltransferase

Query Match 57.8% Score 37; DB 2; Length 505;
Best Local Similarity 54.5% Pred. No. 50;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KGTQGYTDQIE 12
Db 65 KGTQGYOSKLE 75

RESULT 11
S45631
DNA primase chain p58 - human

C/Species: Homo sapiens (man)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C/Accession: S45631

R/Stadlbauer, F.; Bruckner, A.; Rehfuss, C.; Eckerskorn, C.; Lottspeich, F.; Foerster,
Eur. J. Biochem. 222, 781-793, 1994

A/Title: DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.
A/Reference number: S45628; MUID:94298818; PMID:8026492

A/Accession: S45631

A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-509 <STA>

A/Cross-references: EMBL:X74331; NID:G510407; PIDN:CAA52378.1; PID:G510408

C/Supersfamily: mouse DNA primase 54K chain

Query Match 57.8% Score 37; DB 2; Length 509;
Best Local Similarity 54.5% Pred. No. 50;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KGTQGYTDQIE 12
Db 65 KGTQGYOSKLE 75

RESULT 12

S35307
polypotein pp220 precursor - African swine fever virus
N:Contains: structural protein p14; structural protein p150; structural protein p34; str
C:Species: African swine fever virus, ASFV
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #ext_change 24-Nov-1999
C:Accession: S35307, S39809
R:Simon-Mateo, C.; Andres, G.; Vinuela, E.
EMBO J. 12, 2977-2987, 1993
A:Title: Polypotein processing in African swine fever virus: a novel gene expression st
A:Reference number: S35307, MUID:93327788; PMID:8335009
A:Accession: S35307
A:Molecule type: DNA
A:Residues: 1-2475 <SIM>
A:Cross-references: EMBL:222777; NID:G394708; PIDN:CAA80455.1; PID:G394709
A:Accession: S39809
A:Molecule type: protein
A:Residues: 45-51;523-549;894-908 <SI2>
C:Superfamily: African swine fever virus polypotein pp220
C:Keywords: myristylation; polypotein
F:45-522/Product: structural protein p34 #status predicted <MAT1>
F:3-893/Product: structural protein p37 #status predicted <MAT2>
Query Match
Best Local Similarity 57.8%; Score 37; DB 2; Length 2475;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Oy 1 NKGTQYTDQIE 12
Db 395 NKSTQAYNDFLE 406

RESULT 13
T02202
2-hydroxyisoflavone reductase (EC 1.3.1.45) - common tobacco
N:Alternate names: protein A622
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #ext_change 21-Jul-2000
C:Accession: T02202
R:Hiib, N.; Higashiguchi, S.; Hashimoto, T.; Yamada, Y.
Plant Cell 6, 723-735, 1994
A:Title: Gene expression in tobacco low-nicotine mutants.
A:Reference number: Z14614, MUID:94312878; PMID:8038607
A:Accession: T02202
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-310 <HIB>
A:Cross-references: EMBL:D28505; NID:G507822; PIDN:BA05866.1; PID:G507823
A:Experimental source: strain Burley 21; tissue-type root
A:Superfamily: 2'-hydroxyisoflavone reductase
A:Keywords: oxidoreductase
Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 310;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 3 GTQGYTDQI 11
Db 87 GGQGFDTQV 95

RESULT 14
E84404
aminopeptidase homolog [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #ext_change 02-Feb-2001
C:Accession: E84404
R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freites, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: E84404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-367 <STO>
A:Cross-references: GB:AE004437; NID:G10581936; PIDN:AAG20601.1; GSPDB:GN00138
C:Genetics:
A:Gene: pep3
Query Match
Best Local Similarity 54.5%; Score 36; DB 2; Length 367;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Oy 1 NKGTQYTDQI 11
Db 287 NRGIQFTDSI 297

RESULT 15
H85512
hypothetical protein yafa [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #ext_change 14-Sep-2001
C:Accession: H85512
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85512
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <STO>
A:Cross-references: GB:AE005174; NID:G12513013; PIDN:AAG54564.1; GSPDB:GN00145; UWGP:203
C:Genetics:
A:Gene: yafa
C:Superfamily: yafa protein
Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 414;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Oy 1 NKGTQYTDQIE 12
Db 399 DKXQERTDWIE 410

RESULT 16
H69452
hypothetical protein AF1625 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 22-Oct-1999
C:Accession: H69452
R:Kleik, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glock, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utlend, T.; Cotton, M.D.; Spriggs, T.; Attich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Weese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: H69452
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-581 <KLB>
A:Cross-references: GB:AE000990; GB:AE000782; NID:G2689313; PIDN:AAB89628.1; PID:G264893
Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 581;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Oy 1 NKGTQYTDQIE 12
Db 11:|||||

Db 565 NRDSQIYTDYLB 576

RESULT 17

G75362

aconitate hydratase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: G75362

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: G75362

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-906 <WHI>

A:Cross-references: GB:AE002013; GB:AE000513; NID:G6459484; PIDN:AAF11276.1; PID:G645949

A:Experimental source: strain R1

A:Gene: DR1720

A:Map position: 1

C:Superfamily: iron-responsive element-binding protein

Query Match 56.2%; Score 36; DB 2; Length 906;

Best Local Similarity 77.8%; Pred. No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTGQYTDQIE 11

Db 490 GAGQYLDQI 498

RESULT 18

152882

autoantigen - human

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000

C:Accession: 152882

R:Bloch, D.B.; Rabkina, D.; Quertemous, T.; Bloch, K.D.

Clin. Immunol. Immunopathol. 72, 380-389, 1994

A:Title: The immunoreactive region in a novel autoantigen contains a nuclear localization

A:Reference number: 152882; MUID:94340813; PMID:7520377

A:Accession: 152882

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1215 <RBS>

A:Cross-references: GB:L26339; NID:G533201; PIDN:AAA21833.1; PID:G533202

C:Superfamily: human autoantigen Ge-1

Query Match 56.2%; Score 36; DB 2; Length 1215;

Best Local Similarity 60.0%; Pred. No. 2.1e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTGQYTDQIE 12

Db 1144 GTGQYTDQIE 1153

RESULT 19

C66579

hypothetical protein [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: C66579

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishitani, F.; Ouchi, K.; Shiba, T.; Ie

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.

A:Reference number: A66491; MUID:20330349; PMID:10871362

A:Accession: C66579

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-84 <STO>

A:Cross-references: GB:BA000008; NID:G8979082; PIDN:BA09917.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Gene: CP10710

Query Match 54.7%; Score 35; DB 2; Length 84;

Best Local Similarity 54.5%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQYTDQIE 12

Db 24 KGVQYLYELE 34

RESULT 20

F72046

conserved hypothetical protein CP0036 [imported] - Chlamydia pneumoniae (strains CWT

N:Alternate names: c666 hypothetical protein

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: F72046; C81620

R:Katman, S.; Mitchell, W.; Marathe, R.; Lamme, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: F72046

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-84 <ARN>

A:Cross-references: GB:AE001652; GB:AE001363; NID:G4376997; PIDN:AA018449.1; PID:G437701

A:Experimental source: strain CWT029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Grimm, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: C81620

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-84 <REA>

A:Cross-references: GB:AE002167; GB:AE002161; NID:G7188971; PIDN:AAF37931.1; PID:G718897

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CP00710; CP0036

Query Match 54.7%; Score 35; DB 2; Length 84;

Best Local Similarity 54.5%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQYTDQIE 12

Db 24 KGVQYLYELE 34

RESULT 21

H90786

hypothetical protein Ecs1264 [imported] - Escherichia coli (strain O157:H7, substrain R1

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: H90786

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasaara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno-

A:Reference number: A99629; MUID:21156231; PMID:11558796

A:Accession: H90786

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA034667.1; PID:G13360724; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:
A:Gene: ECol1264

Query Match

Best Local Similarity 54.7%; Score 35; DB 2; Length 375;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQGYTDQI 11
DB 230 KGMDDYADQL 239

RESULT 22

H5646

hypothetical protein ycd0 [imported] - Escherichia coli (strain O157:H7, substrain EDL93)

C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H5646

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

R:Perna, L.; Grobbeck, E.J.; Davis, N.W.; Llim, A.; Dimalanta, E.; Potamousta, K.; Apodaca,

Re 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: AB5480; MUID:21074935; PMID:11206551

A:Accession: H5646

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <STO>

A:Cross-references: GB:AE005174; NID:912514380; PIDN:AA655636.1; GSPDB:GN00145; UWGP:215

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: ycd0

Query Match

Best Local Similarity 54.7%; Score 35; DB 2; Length 375;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQGYTDQI 11
DB 230 KGMDDYADQL 239

RESULT 23

Q0EC49

yafa protein, 49k - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 04-Dec-1986 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002

C:Accession: H64748; A04451

R:Plattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

Rose, D.J.; Mau, B.; Shao, Y.

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H64748

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-414 <BLAT>

A:Cross-references: GB:AE000132; GB:U00096; NID:92367098; PIDN:AC73343.1; PID:91786434;

A:Experimental source: strain K-12, substrain MG1655

R:Nuesch, J.; Schumperli, D.

Gene 32, 243-249, 1984

A:Title: Structural and functional organization of the gpt gene region of Escherichia co

A:Accession: A04451

A:Molecule type: DNA

A:Residues: 1-241, 'S', 243-276, 'TL', 279-396, 'KILTKVFRKSPGS', 411, 'NA', 414, 'VKNLLNPNLVKQL

A:Cross-references: GB:M13422; NID:9146255; PIDN:AA23929.1; PID:9146257

C:Comment: The gene coding for this protein is located between the gpt and phoB genes.

C:Genetics:

A:Gene: yafa

A:Map position: 6 min

C:Superfamily: yafa protein

54.7%; Score 35; DB 1; Length 414;

Best Local Similarity 58.3%; Pred. No. 97;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQGYTDQIE 12
DB 399 DKGLQETDWIR 410

RESULT 24

B90662

hypothetical protein ECol266 [imported] - Escherichia coli (strain O157:H7, substrain RI

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: B90662

R:Havashi, T.; Makino, K.; Onishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gaawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: B90662

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-414 <HAV>

A:Cross-references: GB:BA000007; PIDN:BA33689.1; PID:913359723; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: BC60266

C:Superfamily: yafa protein

Query Match

Best Local Similarity 54.7%; Score 35; DB 2; Length 414;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQGYTDQIE 12
DB 399 DKGLQETDWIR 410

RESULT 25

JC7102

histone deacetylase (EC 3.5.1.-) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C:Accession: JC7102

R:Mahnke, U.; Hoelzer, D.; Bucala, R.; Verdin, E.

Biochem. Biophys. Res. Commun. 263, 482-490, 1999

A:Title: Cloning and characterization of the murine histone deacetylase (HDAC3).

A:Reference number: JC7102; MUID:99423490; PMID:10491319

A:Accession: JC7102

A:Molecule type: mRNA

A:Residues: 1-424 <MAH>

A:Cross-references: GB:AF074881

C:Genetics:

A:Gene: hdac3

A:Map position: 18B3

A:introns: 19/1; 46/3; 94/2; 121/3; 140/3; 159/2; 204/1; 231/1; 252/3; 277/2; 307/2; 327

C:Superfamily: RPD protein; RPD3/acuc homology

C:Keywords: hydrolase; phosphoprotein

F/7-307/Domain: RPD3/acuc homology <RAH1>

F/232,282,404/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match

Best Local Similarity 54.7%; Score 35; DB 2; Length 424;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQGYTDQI 11
DB 348 NONSRQYTDQI 358

RESULT 26

JC5834

histone deacetylase (EC 3.5.1.-) 3 - human

C/Species: Homo sapiens (man)
 C/Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 12-Nov-1999
 C/Accession: J05834
 R/Dangond, F.; Haefl, D.A.; Tong, J.K.; Randall, J.; Kojima, R.; Utku, N.; Gullans, S.R.
 Biochem. Biophys. Res. Commun. 242, 648-652, 1998
 A/Title: Differential display cloning of a novel human histone deacetylase (HDAC3) cDNA
 A/Reference number: J05834; MUID:98125547; PMID:9464271
 A/Accession: J05834
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-428 <DAN>
 A/Cross-references: GB:U06914; NID:92934900; PIDN:AA052038.1; PID:g2226173
 C/Comment: This enzyme plays a role in multiple and complex cellular pathways of immune
 gets for the treatment of cancer and autoimmunity.
 C/Superfamily: RPD3 protein; RPD3/acuc homology
 C/Keywords: hydrolase; phosphoprotein
 F.7.3.07/Domain: RPD3/acuc homology <RAH1>
 F.7.3.282,404/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 54.7%; Score 35; DB 2; Length 428;
 Best Local Similarity 54.5%; Pred. No. 1e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQGYTDQIE 11
 | : : : | : : : |
 Db 348 NNSRQYLDQI 358

RESULT 27
 F82070
 2-isopropylmalate synthase VC2490 [!imported] - Vibrio cholerae (strain N16961 serogroup
 C/Species: Vibrio cholerae
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C/Accession: F82070
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 Charsorn, D.; Esmolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dregoi, I.; Sellers, F.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A/Reference number: A82035; MUID:20406833; PMID:10952301
 A/Accession: F82070
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-516 <HEI>
 A/Cross-references: GB:AE004318; GB:AE003852; NID:99657063; PIDN:AA05632.1; GSPDB:GN001
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Position: VC2490
 A/Position: 1
 C/Superfamily: 2-isopropylmalate synthase leuA

Query Match 54.7%; Score 35; DB 2; Length 516;
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQGYTDQIE 12
 | : : : | : : : |
 Db 127 KHARYTDQVE 137

RESULT 28
 G87586
 cation transporting ATPase CC2726 [!imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 23-Dec-2002
 C/Accession: G87586
 R/Nietman, W.C.; Faldutym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon
 n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of Caulobacter crescentus.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: G87586

A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-677 <STO>
 A/Cross-references: GB:AE005673; NID:g13424315; PIDN:AAK24691.1; GSPDB:GN00148
 C/Genetics:
 A/Accession: CC2726
 A/Gene: CC2726
 C/Superfamily: probable cadmium-transporting ATPase P603.150; ATPase nucleotide-binding

Query Match 54.7%; Score 35; DB 2; Length 677;
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 TQGYTDQIE 12
 | : : : | : : : |
 Db 273 TQGYTDQIE 281

RESULT 29
 S76082
 hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C/Species: Synechocystis sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 02-Aug-2002
 C/Accession: S76082
 R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimizu, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-116, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 sp.
 A/Reference number: S74322; MUID:97061201; PMID:8905231
 A/Accession: S76082
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-714 <KAN>
 A/Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BA10060.1; PID:dt01071
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C/Genetics:
 A/Start codon: GTG
 C/Superfamily: Nostoc sp. hypothetical protein al12707; dnaJ amino-terminal homology
 F.6-70/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 54.7%; Score 35; DB 2; Length 714;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTQGYTDQIE 12
 | : : : | : : : |
 Db 386 NPSVOGYLDQIE 397

RESULT 30
 T22615
 hypothetical protein T01B8.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T22615; T24303
 R/McMurray, A.
 submitted to the EMBL Data Library, March 1995
 A/Reference number: T22615
 A/Accession: T22615
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-1270 <WIL>
 A/Cross-references: EMBL:Z48583; PIDN:CAA86473.1; GSPDB:GN000020; CESP:T01B8.5
 R/McMurray, A.
 submitted to the EMBL Data Library, March 1995
 A/Reference number: Z19871
 A/Accession: T24303
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-1270 <WIL>
 A/Cross-references: EMBL:Z48809; PIDN:CAA86749.1; GSPDB:GN000020; CESP:T01B8.5

A:Experimental source: clone T01E8
C:Genetics:
A:Gene: CESP:T01E8.5
A:Map position: 2
A:Insertion: 15/1, 35/3; 73/3; 99/2, 237/1; 277/3; 332/3; 397/2; 534/3; 703/3; 1006/1; 115

Query Match 53.3%; Score 34.5; DB 2; Length 1270;
Best Local Similarity 53.8%; Pred. No. 4.2e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 NKCT-QOYTDQIE 12
DB 1038 NRGTYKRCQDID 1050

RESULT 31
A89872
Conserved hypothetical protein SA0887 [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
Accession: A89872

med. A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;
C.; Shida, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11416146

A:Accession: A89872

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-106 <R>

A:Cross-references: GB:BA000018; PID:G13700836; PIDN:BA842132.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SN0887

Query Match 53.1%; Score 34; DB 2; Length 106;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 45 KGTQOYTN 52

RESULT 32

D69968
Hypothetical protein yqxJ - Bacillus subtilis

C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
Accession: D69968; E44816

R.Kunze, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Ertan, K.D.; Ertan, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulter, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetler, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidos, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maesl
R.; M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetle
Yeger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A.; Authors: Schleicher, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akuch, M.; Tanakoshi, A.; Tanaka, T.; Tetsuya, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; PMID:9804033; PMID:9384377

A:Accession: D69968

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-120 <R>

A:Cross-references: GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CAB14529.1; PID:G2635033

A:Experimental source: strain 168

R:Foster, S.J.
J. Gen. Microbiol. 137, 1987-1998, 1991

A:Title: Cloning, expression, sequence analysis and biochemical characterization of an a
A:Reference number: A44816; PMID:92065258; PMID:1683402

A:Accession: E44816

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'L', 2-120 <FOS>

A:Note: sequence extracted from NCBI backbone (NCBIN:66026, NCBI:66043)

C:Genetics:

A:Gene: yqxJ

C:Superfamily: Bacillus subtilis hypothetical protein yqxJ

Query Match 53.1%; Score 34; DB 2; Length 120;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 TQOYTDQIE 12
DB 58 TQOYTDKIK 66

RESULT 33

H82553

3-demethylubiquinone-9-3-methyltransferase XP2471 [Imported] - Xylella fastidiosa (strain
C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

A:Accession: H82553

C:Authors: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; PMID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: H82553

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-246 <SIM>

A:Cross-references: GB:AE004055; GB:AE003849; NID:99107661; PIDN:AAPE5269.1; GSPDB:GN001

A:Experimental source: strain 945c

R.Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Birones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmiter, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tauhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2471

C:Superfamily: 3-demethylubiquinone-9-3-O-methyltransferase; bioc homology

Query Match 53.1%; Score 34; DB 2; Length 246;
Best Local Similarity 70.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQOYTDQI 11
DB 182 KGTQOYKDFI 191

RESULT 34
H59128
Conserved hypothetical protein MTH232 - Methanobacterium thermoautotrophicum (strain Del

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999

A:Accession: H59128

R:Smith, D.R.; Doucet-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N. K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct
 A|Reference number: A69000; MUID:98037514; PMID:9371463
 A|Accession: H69128
 A|Status: preliminary; nucleic acid sequence not shown; translation not shown
 A|Molecule type: DNA
 A|Residues: 1-255 <MT>
 A|Cross-references: GB:AE000810; GB:AE000666; NID:g2621277; PIDN:AAB84738.1; PID:g262128
 A|Experimental source: strain Delta H
 C|Genetics:
 A|Gene: MTH232
 C|Superfamily: conserved hypothetical protein YBR0202c

Query Match 53.1%; Score 34; DB 2; Length 255;
 Best Local Similarity 50.0%; Pred. No. 87;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

2 KGTQOYTDQI 11
 |||||:
 DB 139 KSTEGYSDRL 148

RESULT 35
 T03635
 plastid-lipid-associated protein pap - common tobacco (fragment)
 N|Alternate names: probable fibrillin
 C|Species: Nicotiana tabacum (common tobacco)
 C|Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
 C|Accession: T03635
 R|Pozueta-Romero, J.; Rafia, F.; Houline, G.; Chéniclet, C.; Carde, J.P.; Schantz, M.L.;
 Plant Physiol. 115, 1185-1194, 1997
 A>Title: A ubiquitous plant housekeeping gene, PAP, encodes a major protein component of
 A|Reference number: Z14983; MUID:98052146; PMID:9390444
 A|Accession: T03635
 A|Status: preliminary; translated from GB/EMBL/DBJ
 A|Molecule type: DNA
 A|Residues: 1-270 <PO>
 A|Cross-references: EMBL:Y15489; NID:g2632087; PIDN:CAA75657.1; PID:g2632088
 A|Experimental source: strain xanthi NC; tissue-type leaves
 C|Genetics:
 A|Note: pap
 A|Note: Intron positions not resolved (incomplete sequence)
 C|Keywords: chromoplast

Query Match 53.1%; Score 34; DB 2; Length 270;
 Best Local Similarity 70.0%; Pred. No. 93;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

3 GTOQYTDQIE 12
 |||||:
 DB 175 GTPQUTDSIE 184

RESULT 36
 B89991
 truncated beta-hemplysin [imported] - *Staphylococcus aureus* (strain N315)
 C|Species: *Staphylococcus aureus*
 C|Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C|Accession: B89991
 R|Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
 ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A|Reference number: A89758; MUID:21311952; PMID:11418146
 A|Accession: B89991
 A|Status: preliminary
 A|Molecule type: DNA
 A|Residues: 1-289 <KUR>
 A|Cross-references: GB:BA000018; PID:g13701798; PIDN:BA043091.1; GSPDB:GN00149
 A|Experimental source: strain N315

C|Genetics:
 A|Gene: truncated(hnb)

Query Match 53.1%; Score 34; DB 2; Length 289;
 Best Local Similarity 41.7%; Pred. No. 1e+02;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 |||||:
 DB 190 NKGTPEPDMUK 201

RESULT 37
 AG3278
 hypothetical protein BME10212 [imported] - *Brucella melitensis* (strain 16M)
 C|Species: *Brucella melitensis*
 C|Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C|Accession: AG3278
 R|DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A>Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A|Reference number: AD3252; PMID:11756688
 A|Accession: AG3278
 A|Status: preliminary
 A|Molecule type: DNA
 A|Residues: 1-299 <KUR>
 A|Cross-references: GB:AE006917; PIDN:AAL51394.1; PID:g17982098; GSPDB:GN00190
 A|Experimental source: strain 16M
 C|Genetics:
 A|Gene: BME10212
 A|Map position: 1

Query Match 53.1%; Score 34; DB 2; Length 299;
 Best Local Similarity 54.5%; Pred. No. 1e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQI 11
 |||||:
 DB 165 NKGTGVDRI 175

RESULT 38
 H75378
 hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C|Species: *Deinococcus radiodurans*
 C|Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C|Accession: H75378
 R|White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Shen, H.O.; Vamathevan, J.T.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A|Reference number: A75250; MUID:20036896; PMID:10567266
 A|Accession: H75378
 A|Status: preliminary
 A|Molecule type: DNA
 A|Residues: 1-304 <WH>
 A|Cross-references: GB:AE002002; GB:AE000513; NID:g6459345; PIDN:AAFI1154.1; PID:g645935
 A|Experimental source: strain R1
 C|Genetics:
 A|Gene: DR1591
 A|Map position: 1
 C|Superfamily: *Deinococcus radiodurans* hypothetical protein DR1591

Query Match 53.1%; Score 34; DB 2; Length 304;
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 |||||:
 DB 286 NKGLNEYVSIE 297

RESULT 39

T04905
 fibrillin homolog T10114.70 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
 C:Accession: T04905
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15389
 A:Accession: T04905
 A:Molecule type: DNA
 A:Residues: 1-310 <BBV>
 A:Cross-references: EMBL:AL021712
 A:Experimental source: cultivar Columbia; BAC clone T10114
 C:Genetics:
 A:Map position: 4
 A:Introns: 141/2; 206/3
 A:Note: T10114.70

Query Match

53.1%; Score 34; DB 2; Length 310;
 Best Local Similarity 70.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GTQGYTDQIE 12
 ||| |||
 DB 215 GTPQLTDSIE 224

RESULT 40

T01472
 Probable fibrillin T24H24.16 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Mar-1999
 C:Accession: T01472
 R:Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K.
 submitted to the EMBL Data Library, August 1998
 A:Description: The sequence of A. thaliana T24H24.
 A:Reference number: Z14333
 A:Accession: T01472
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-318 <COU>
 A:Cross-references: EMBL:AF075598; NID:G3293581; PID:G3377825
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Map position: 4
 A:Introns: 149/2; 214/3
 A:Note: T24H24.16

Query Match

53.1%; Score 34; DB 2; Length 318;
 Best Local Similarity 70.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GTQGYTDQIE 12
 ||| |||
 DB 223 GTPQLTDSIE 232

Search completed: August 20, 2003, 09:32:37
 Job time : 15.4845 secs


```

DE 60 kDa chaperonin (Protein Cpn60) (groEL protein) .
GN GROL OR GROEL OR MOA OR PM1107.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=96105224; PubMed=8529887;
RA Love B.C., Hansen L.M., Hirsch D.C.;
RT "Cloning and sequence of the groEL heat-shock operon of Pasteurella
RL multocida.";
RN Gene 166:179-180(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RA Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
-1- FUNCTION: Prevents misfolding and promotes the refolding and
proper assembly of unfolded polypeptides generated under stress
conditions (By similarity).
-1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
7 subunits (By similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U30165, AAA84916.1, -.
DR EMBL, AE006151, AAK03191.1, -.
DR PIR, JC4519, JC4519.
DR HSSP, P06139, IJON.
DR HAMAP, MF_006000, -, 1.
DR InterPro, IPR001844, Chaperon Cpn60.
DR InterPro, IPR002423, Cpn60/TCF-1.
DR Pfam, PF00118, Cpn60_TCF1, 1.
DR PRINTS, PR00298, CHAPERONIN60.
DR PRINTS, PR00304, TCOMPLEXTCF1.
DR PROSITE, PS00296, CHAPERONIN60, 1.
DR Chaperone; ATP-binding; Complete proteome.
DR CONFLICT 424 424 S -> N (IN REF. 1).
DR CONFLICT 464 464 I -> V (IN REF. 1).
FT FT SEQUENCE 547 AA; 57291 MW; EBB95B9F2158B55E CRC64;
SQ
Query Match 67.2%; Score 43; DB 1; Length 547;
Best Local Similarity 66.7%; Pred. No. 2.3;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0.
QY 1 NKGTOQYTDQIE 12
Db 479 NAGTQYGDME 490

```

```

OC Pasteurellaceae; Haemophilus.
CX NCBI_TaxID=727;

RN
RP STRAIN=RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White
RA Kesteven A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Scott J.D., Shiley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uetereback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Gnehm L.D., Fitchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA Grahe C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD Rd."
RL
CC Science 269:496-512(1995).
CC -I- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -I- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or\_send\_an\_email\_to\_license@isb-sib.ch).
CC -----
CC EMBL; U32736; AAC22201.1; -.
DR PIR; G64076; G64076.
DR HSPD; P06139; IJON.
DR TIGR; H10543; -.
DR HAMAP; MF_00600; -; 1.
DR InterPro; IPR001844; Chaperonin Cpn60.
DR InterPro; IPR002423; Cpn60/TCE-1.
DR Pfam; PF00118; Cpn60 TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
DR Chaperone; ATP-binding; Complete proteome.
CX SEQUENCE 548 AA; 57577 MW; CA4066AAC1B62159 CRC64;

Query Match 67.2%; Score 43; DB 1; Length 548;
Beet Local Similarity 66.7%; Pred. No. 2.3;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 NKGQOYTDPLE 12
| | | | | | |
Db 480 NAGTEGYDWMTE 491

RESULT 4
CH60_ACTPL STANDARD; PRT; 546 AA.
AC P9416;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROU OR GROEL OR MOXA.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
CX NCBI_TaxID=715;
RN
RP SEQUENCE FROM N.A.
RX STRAIN=S 4074 / Serotype 1;

```



```

RX MEDLINE=97189570; PubMed=9037757;
RA Vezina G., Stros M., Clairoux N., Boissinot M.;
RT "Cloning and characterization of the groS locus from Actinobacillus
RT pleuropneumoniae.";
RL FEWS Microbiol. Lett. 147:11-16(1997).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (Hsp60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: U55016; AAB51437.1; -.
DR HSSP; P06139; IGR1.
DR HAMAP; MF_00600; -.
DR InterPro; IPR001844; Chaperin_Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONIN_CPN60; 1.
DR Chaperone; ATP-binding.
KW Chaperone; ATP-binding.
FT INIT MET 0
SQ SEQUENCE 546 AA; 57513 MW; E84B72C9BD3DB56 CRC64;

Query Match 64.1%; Score 41; DB 1; Length 546;
Best Local Similarity 58.3%; Pred. No. 5.3;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 NKGQOYTDQI 12
Db 479 NAGTEQYDQMI 490

RESULT 5
CH60_ACTAC STANDARD; PRT; 546 AA.
AC P46398;
01-NOV-1995 (Rel. 32, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR GROEL OR MOXA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
CX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39.
RC STRAIN=Y4;
RX MEDLINE=96017061; PubMed=7567064;
RA Nakano T., Inai Y., Yamashita Y., Kusuzaki-Nagira T., Nagaoka S.,
RA Okahashi N., Koga T., Nishihara T.;
RT "Molecular and immunological characterization of a 64-kDa protein of
RT Actinobacillus actinomycetemcomitans.";
RL Oral Microbiol. Immunol. 10:151-159(1995).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the chaperonin (Hsp60) family.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: D28817; BAB05977.1; -.
DR HSSP; P06139; IGR1.
DR HAMAP; MF_00600; -.
DR InterPro; IPR001844; Chaperin_Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONIN_CPN60; 1.
DR Chaperone; ATP-binding.
KW Chaperone; ATP-binding.
FT INIT MET 0
SQ SEQUENCE 546 AA; 57312 MW; 6249E0E46334F6F6 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 546;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 NKGQOYTDQI 11
Db 478 NAGTEQYDQMI 488

RESULT 6
G7AC_BREDI STANDARD; PRT; 720 AA.
AC Q915D6;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Glutaryl-7-aminocaprolactam precursor (EC 3.5.1.-)
DE (Glutaryl-7-ACA-acylase).
OS Brevundimonas diminuta (Pseudomonas diminuta).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Brevundimonas.
CX NCBI_TaxID=293;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAC-1;
RA Kim D.-W., Kang S.-M., Yoon K.-H.;
RT "Cloning and the nucleotide sequence of a Pseudomonas diminuta KAC-1
RT glutaryl-7-aminocaprolactam precursor acid acylase gene.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- PATHWAY: SECOND STEP IN THE BIOCONVERSION OF CEPHALOSPORIN TO 7-
CC ACA.
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S45.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: AF251710; AAF64242.1; -.
DR PDB; 1FM2; 15-AUG-01.
DR PDB; 1JWZ; 01-SEP-02.
DR PDB; 1JW0; 01-SEP-02.
DR PDB; 1KEH; 04-DEC-02.
DR MEROPS; S45.002; -.
DR InterPro; IPR002692; Penicill_amidase.
DR InterPro; IPR01804; Penicill_amidase; 1.
DR Pfam; PF01804; Penicill_amidase; 1.
KW Hydrolase; Penicillamc; Antibiotic resistance; Zymogen; Signal;

```

KW 3D-structure. 1 27 POTENTIAL.
 FT SIGNAL 28 720 GLUTARYL 7-AMINOCEPHALOSPORANIC ACID
 FT CHAIN 28 720 ACYLASE
 FT ACT SITE 199 199 PROBABLE.
 SQ SEQUENCE 720 AA; 79779 MW; AD624797845CC39B CRC64;
 Query Match 59.4%; Score 38; DB 1; Length 720;
 Best Local Similarity 70.0%; Pred. No. 26;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTGQYTDQIE 12
 DB 680 GTTHYSQDIE 689

RESULT 7
 BIOD_VIBPA STANDARD; PRT; 228 AA.
 ID BIOD_VIBPA STANDARD; PRT; 228 AA.
 0870N3;
 15-SEP-2003 (Rel. 42, Created)
 15-SEP-2003 (Rel. 42, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 Dethiobiotin synthetase (EC 6.3.3.3) (Dethiobiotin synthase) (DTB synthetase) (DTBS).
 GN BIOD OR VP116.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrrio.
 CC NCBI_TaxID=670;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K., Yasunaga T., Honda T., Shinagawa H., Kubota Y., Kimura S., Yajima Y., Nishida M., Nakano M., Yamashita A., Kubota Y., Kimura S., Yajima Y., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae.";
 RL Lancet 361:743-749(2003).
 CC -1- CATALYTIC ACTIVITY: ATP + 7,8-diaminononanoate + CO(2) = ADP + phosphate + dethiobiotin.
 CC -1- COFACTOR: Magnesium (By similarity).
 CC -1- PATHWAY: Bioconversion of pimelate into dethiobiotin.
 CC -1- SIMILARITY: BELONGS TO THE DETHIOBIOTIN SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF005076; BACS9379.1; ALT_INIT.
 DR HAMAP; MF_00336; -; 1.
 KW Biotin biosynthesis; Ligase; Magnesium; ATP-binding;
 KM Complete proteome.
 FT NP_BIND 9 17 ATP (BY SIMILARITY).
 SQ SEQUENCE 228 AA; 24711 MW; 9B10E29B7B42DF6 CRC64;
 Query Match 57.8%; Score 37; DB 1; Length 228;
 Best Local Similarity 58.3%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 DB 179 NPGTEHYADIE 190

RESULT 8
 TRY3_SALSA STANDARD; PRT; 238 AA.
 ID TRY3_SALSA STANDARD; PRT; 238 AA.

AC P35033;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Trypsin III precursor (EC 3.4.21.4) (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Procanthopterygii; Salmoniformes; Salmonidae; Salmo.
 CC NCBI_TaxID=8030;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=96035908; PubMed=7556223;
 RA Male R., Lorens J.B., Smales A.O., Torrissen K.R.;
 RT "Molecular cloning and characterization of anionic and cationic variants of trypsin from Atlantic salmon.";
 RL Eur. J. Biochem. 232:677-685(1995).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO TRYPSIN FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70074; CAA49679.1; -;
 DR PIR; S66657; S31779.
 DR PDB; 1A0J; 13-JAN-99.
 DR MEROPS; S01.151; -;
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; Trypsin; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00340; TRYPSIN_HIS; 1.
 DR PROSITE; PS00335; TRYPSIN_SER; 1.
 KM Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family; 3D-structure.
 FT FT NON_TER 1 7
 FT SIGNAL 1 7
 FT PROPEP 8 15
 FT CHAIN 16 238
 FT ACT_SITE 55 55
 FT ACT_SITE 99 99
 FT ACT_SITE 192 192
 FT DISULFID 22 152
 FT DISULFID 40 56
 FT DISULFID 124 225
 FT DISULFID 131 198
 FT DISULFID 163 177
 FT DISULFID 188 212
 FT SITE 186 186
 SQ SEQUENCE 238 AA; 25389 MW; AE799B808393023 CRC64;
 Query Match 57.8%; Score 37; DB 1; Length 238;
 Best Local Similarity 41.7%; Pred. No. 12;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 DB 73 NEGTEQFDISVK 84

RESULT 9
 PRI2_MOUSE STANDARD; PRT; 505 AA.
 ID PRI2_MOUSE STANDARD; PRT; 505 AA.
 AC P33610;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)

```

DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DNA primase large subunit (EC 2.7.7.-) (DNA primase 58 kDa subunit)
GN PRIM2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-48; 186-212 AND 403-416.
RX MEDLINE=93216788; PubMed=8463324;
RA Miyazawa H., Izumi W., Tada S., Takada R., Masutani M., Ui M.,
RA Hanaoka F.;
RT "Molecular cloning of the cDNAs for the four subunits of mouse DNA
RT polymerase alpha-primase complex and their gene expression during
RT cell proliferation and the cell cycle."
RL J. Biol. Chem. 268:8111-8122 (1993).
RN [2]
RP SEQUENCE FROM N.A.
R Brueckner A., Stadlbauer F., Tseng B.Y., Naehner H.P.;
Submitted (Aug-1993) to the EMBL/Genbank/DBJ databases.
-1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
DNA REPLICATION.
-1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
-1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE LARGE SUBUNIT
FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: D13545; BAA02745.1; -
CC EMBL: D17385; BAA04203.1; -
CC PIR: C46642; C46642.
CC MGD: MGI:97758; Prim2.
DR Pfam: PF04104; DNA_primase_lrg.1.
DR Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome;
KW DNA-binding.
FT CONFLICT 109 109 A -> G (IN REF. 2).
SQ SEQUENCE 505 AA; 58408 MW; 36A64C77669D8C24 CRC64;

Query Match
Best Local Similarity 57.8%; Score 37; DB 1; Length 505;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KGTQGYTDQIE 12
DB 65 KGTQGYTSKLE 75

RESULT 10
PR12 HUMAN STANDARD; PRT; 509 AA.
AC P49643;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA primase large subunit (EC 2.7.7.-) (DNA primase 58 kDa subunit)
DE (p58).
GN PRIM2A OR PRIM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94298818; PubMed=8026492;
R Stadlbauer F., Brueckner A., Rehness C., Eckerskorn C.,

```

```

RA Lottspeich F., Foerster V., Tseng B.Y., Naehner H.P.;
RT "DNA replication in vitro by recombinant DNA-polymerase-alpha-
RT primase."
RL Eur. J. Biochem. 222:781-793 (1994).
CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
CC DNA REPLICATION.
CC -1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE LARGE SUBUNIT
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: X74331; CAA52378.1; -
CC PIR: S45631; S45631.
DR Genem: HGNC:9370; PRIM2A.
DR GK: P49643; -
DR MIM: 176636; -
DR GO: GO:0003896; P:DNA primase activity; TAS.
DR GO: GO:0006269; P:DNA replication, priming; TAS.
DR Pfam: PF04104; DNA_primase_lrg.1.
KW Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome;
KW DNA-binding.
SQ SEQUENCE 509 AA; 58778 MW; B1F500EC442BEF9C CRC64;

Query Match
Best Local Similarity 57.8%; Score 37; DB 1; Length 509;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 KGTQGYTDQIE 12
DB 65 KGTQGYTSKLE 75

RESULT 11
MSS_CYPCA STANDARD; PRT; 1935 AA.
AC Q90339;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthys; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=97352533; PubMed=9208928;
RA Hirayama Y., Watabe S.;
RT "Structural differences in the crossbridge head of temperature-
RT associated myosin subfragment-1 isoforms from carp fast skeletal
RT muscle."
RL Eur. J. Biochem. 246:380-387 (1997).
RN [2]
RP SEQUENCE OF 981-1935 FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=97176447; PubMed=9023993;
RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;
RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal
RT muscle and their gene expression associated with temperature
RT acclimation."
RL J. Exp. Biol. 200:27-34 (1997).
RN [3]
RP SEQUENCE OF 1387-1528 FROM N.A.

```

RX MEDLINE=95194396; PubMed=7887920;
 RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H.,
 RA Uozumi T., Hirono I., Aoki T.;
 RT "Temperature acclimation induces light meromyosin isoforms with
 RT different primary structures in carp fast skeletal muscle.";
 RL Biochem. Biophys. Res. Commun. 208:118-125(1995).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC).
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D89992; BAA22069.1; -
 CC EMBL; D50476; BAA09069.1; -
 CC EMBL; D43700; BAA07802.1; -
 CC PIR; I50496; I50496.
 CC HSR; P13538; 2MTS.
 CC InterPro; IPR000048; IQ_region.
 CC InterPro; IPR001609; myosin_head.
 CC InterPro; IPR004009; Myosin_N.
 CC InterPro; IPR002928; Myosin_tail.
 CC Pfam; PF00612; IQ_1.
 CC Pfam; PF00063; myosin_head; 1.
 CC Pfam; PF02736; Myosin_N; 1.
 CC Pfam; PF01576; Myosin_tail; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC ProDom; PD000355; myosin_head; 1.
 CC SMART; SM00015; IQ; 1.
 CC SMART; SM00242; MISC; 1.
 CC PROSITE; PS50096; IQ; 1.
 CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 CC ATP-binding; Methylation; Alkylation; Calmodulin-binding;
 CC Multigene family.
 KW DOMAIN 1 781 MYOSIN HEAD-LIKE.
 FT DOMAIN 1 782 IQ.
 FT DOMAIN 812 839 HINGE.
 FT DOMAIN 840 1935 COILED COIL (POTENTIAL).
 FT NP_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 659 681 ACTIN-BINDING.
 FT DOMAIN 761 775 ACTIN-BINDING.
 FT MOD_RES 129 129 METHYLATION (TR1-) (POTENTIAL).
 FT MOD_RES 696 696 ALKYLATION (SH-2).
 FT MOD_RES 706 706 ALKYLATION (SH-2).
 FT MOD_RES 706 706 ALKYLATION (SH-2).
 SQ SEQUENCE 1935 AA; 22159 MW; 9A1244B67D65C8B CRC64;

Query Match 57.8%; Score 37; DB 1; Length 1935;
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KGTQYTDQI 12
 DB 1304 RKGQVATQI 1314

RESULT 12
 ID IFRH_TOBAC STANDARD; PRT; 310 AA.
 AC P52579;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Isoflavone reductase homolog A622 (EC 1.3.1.-).
 OS Nicotiana tabacum (Common tobacco).
 CC Bacteria; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
 CC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Burley 21; TISSUE=Root;
 RX MEDLINE=94312878; PubMed=8038607;
 RA Hibl N., Higaehiguchi S., Hashimoto T., Yamada Y.;
 RT "Gene expression in tobacco low-nicotine mutants.";
 RL Plant Cell 6:723-735(1994).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: STRONG, TO ISOFLAVONE REDUCTASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D28505; BAA05866.1; -
 CC PIR; T02202; T02202.
 CC InterPro; IPR003866; Isoflav_reduct.
 CC Pfam; PF02716; Isoflavone_red; 1.
 KW Oxidoreductase; NADP.
 SQ SEQUENCE 310 AA; 34654 MW; 03A7F3B6D451885 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 310;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGTQYTDQI 11
 DB 87 GGTQYTDQV 95

RESULT 13
 ID YG25_ARCFU STANDARD; PRT; 581 AA.
 AC 028648;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1625 precursor.
 GN AF1625.
 OS Archaeoglobus fulgidus.
 CC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 CC Archaeoglobaceae; Archaeoglobus.
 CC NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kexlavage A.R., Graham D.E., Kyriakides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Goehne U.D., Weidman J.F., McDonald L., Uutterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AE000990; AAB89628.1; -
 DR PIR; H69452; H69452.
 DR TIGR; AF1625; -
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 28
 CHAIN 1 28
 SEQUENCE 581 AA; 64858 MW; EABA65AF624EB75C CRC64;
 Query Match 56.2%; Score 36; DB 1; Length 581;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 NKGOQYTDQIE 12
 Db 565 NRDSQYTDYIE 576
 RESULT 14
 Y710 CHLPN STANDARD; PRT; 84 AA.
 AC Q92705;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein CPN0710/CPN036/CPJ0710.
 GN CPN0710 OR CPN036 OR CPJ0710.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydia; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CM1029;
 RX MEDLINE=9920606; PubMed=10192388;
 KA Kaiman S., Mitchell W., Marathe R., Fan J., Hyman R.W.,
 OL Olinger L., Gilwood J., Davis R.W., Stephens R.S.;
 RL "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RN Nat. Genet. 21:385-389(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RA "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RL from Upan and CHL029 from USA.";
 RN Nucleic Acids Res. 28:2311-2314(2000).
 RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RA Hirakawa H., Shirai M., Kohara S.;
 RT "Genomic sequence comparison of two unrelated isolates of Chlamydia
 RL pneumoniae from Japan and U.S.";
 CC CC Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0710/CT666/TC0037
 CC CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AE001652; AAD18849.1; -
 DR EMBL; AE002167; AAF37931.1; -
 DR EMBL; AP002547; BAA38917.1; -
 DR EMBL; AB035952; BAA86660.1; -
 DR PIR; C86579; C86579.
 DR PIR; F72046; F72046.
 DR PHCI-ZDPAGE; Q92705; -
 DR TIGR; CPN036; -
 KW Complete proteome.
 SQ SEQUENCE 84 AA; 9348 MW; 50A9CC271A5CB693 CRC64;
 Query Match 54.7%; Score 35; DB 1; Length 84;
 Best Local Similarity 54.5%; Pred. No. 9.4;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 KGTOQYTDQIE 12
 Db 24 KGVQYLTLEL 34
 RESULT 15
 YAF4 ECOLI STANDARD; PRT; 414 AA.
 AC P04335; P77413;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yaf4.
 GN YAF4 OR B0239.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=8515481; PubMed=6397401;
 RA Nueesch J., Schumperli D.;
 RT "Structural and functional organization of the gpt gene region of
 RL Escherichia coli.";
 RN Gene 32:243-249(1984).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor U., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;

```

RT      4.0 - 6.0 min (189,987 - 281,456b) region." ;
RT      Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA      Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurd O.,
RA      Lahekari D., Lew H., Lin D., Nemach A., Oefner P., Roberts D.,
RA      Davis R.W.,
RL      Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sb.ch/announce/
CC      or send an email to license@isb-sb.ch).
CC      -----
DB      EMBL; AB000432; AAC73393.1; -.
DB      EMBL; AB000432; AAC73393.1; -.
DB      EMBL; D83536; BAA77908.1; -.
DB      EMBL; U70214; AAB08659.1; -.
DR      PIR; H64748; QOEC49.
DR      Ecogene; EG11091; yafA.
DR      InterPro; IPR000379; Ser_actre_site.
KM      Hypothetical protein; Complete proteome.
FT      CONFLICT    242    242
FT      CONFLICT    277    278
FT      CONFLICT    397    414
FT      FT
FT      FT
SQ      SEQUENCE    414 AA;  47008 MW;  2C154E46CFC965E CRC64;
Query Match
Best Local Similarity    54.7%;    Score 35; DB 1; Length 414;
Matches    7; Conservative    2; Mismatches    3; Indels    0; Gaps    0;
QY      1 NKGTOQYTDQTE 12
DB      399 DKGLQETDWTIE 410
RESULT 16
HDA3_MOUSE
ID      HDA3_MOUSE
AC      O88895; O88896;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
OC      Histone deacetylase 3 (HD3).
ON      HDAC3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RA      STRAIN=C3H, and C57BL/6J;
RA      Mahlknecht U.;
RL      Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      INTERACTION WITH HDAC7.
RX      PubMed=10984530;
RA      Downes M., Ordenllich P., Kao H.-Y., Alvarez J.G.A., Evans R.M.;
RT      "Identification of a nuclear domain with deacetylase activity.";
RC      Proc. Natl. Acad. Sci. U.S.A. 97:10330-10335 (2000).
CC      -I- FUNCTION: Responsible for the deacetylation of lysine residues on
CC      the N-terminal part of the core histones (H2A, H2B, H3 and H4).
CC      Histone deacetylation gives a tag for epigenetic repression and
CC      plays an important role in transcriptional regulation, cell cycle
CC      progression and developmental events. Histone deacetylases act via
CC      the formation of large multiprotein complexes. Required to repress
CC      transcription of the POU1f1 transcription factor (By similarity).

```

```

-1- SUBUNIT: Interacts with HDAC7. Forms a heterologous complex at
CC with with MYI. Interacts with HDAC10 and DAXX. Found in a complex
CC with NCOR1 and NCOR2 (By similarity).
CC
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=O88895-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O88895-2; Sequence=VSP_002080;
CC -1- PTM: Sumoylated in vitro.
CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF074881; AAC36305.1; -
CC EMBL; AF074882; AAC36306.1; -
CC EMBL; AF079310; AAC67258.1; -
CC EMBL; AF079309; AAC67258.1; JOINED.
CC PIR; JCT102; JCT102.
CC MGD; MGI:1343091; Hdac3.
CC InterPro; IPR000286; Hist.deacetylase.
CC Pfam; PF00850; Hist.deacetyl1.
CC PRINTS; PR01270; HDASUPER.
CC
CC Hydroxylase; Nuclear protein; Chromatin regulator;
CC Transcription regulation; Repressor; Ub1 conjugation;
CC Alternative splicing.
CC
CC DOMAIN 3 316 HISTONE DEACETYLASE.
CC ACT SITE 135 135 BY SIMILARITY.
CC VASAPLIC 69 259 Missing (in isoform short).
CC /FtId=VSP_002080.
CC
CC SEQUENCE 424 AA; 48364 MW; B0C3BB2C5CD95E0B CRC64;
CC
CC
CC Query Match 54.7%; Score 35; DB 1; Length 424;
CC Best Local Similarity 54.5%; Pred. No. 53;
CC Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0.
CC
CC Oy 1 NKGTQGYTDQI 11
CC : : : : :
CC
CC Db 348 NQNSROYLDQI 358
CC
CC
CC RESULT 17
CC HDAC3_CHICK STANDARD; PRT; 428 AA.
CC ID HDAC3_CHICK
CC AC P56520;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Histone deacetylase 3 (Hd3).
CC GN HDAC3.
CC OS Gallus gallus (Chicken).
CC CC Eukaryota; Metazoa; Chordata; Vertebrata; Buteleostomi;
CC CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC CC Gallus.
CC CC NCBI_TaxID=9031;
CC CC [1]
CC RN SEQUENCE FROM N.A.
CC RA Takami Y.;
CC RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: Responsible for the deacetylation of lysine residues on
CC the N-terminal part of the core histones (H2A, H2B, H3 and H4).
CC Histone deacetylation gives a tag for epigenetic repression and
CC plays an important role in transcriptional regulation, cell cycle
CC progression and developmental events. Histone deacetylases act via
CC the formation of large multiprotein complexes (By similarity).

```

CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF039753; AAB96925.1;
 CC InterPro: IPR000286; Hist.deacetylase.
 CC Pfam: PF00850; Hist.deacetyl. 1.
 CC PRINTS: PR01270; HDASUPER.
 CC Hydrolyase; Nuclear protein; Chromatin regulator;
 CC Transcription regulation; Repressor.
 CC DOMAIN 3 316 HISTONE DEACETYLASE.
 CC ACT SITE 135 135 BY SIMILARITY.
 CC SEQUENCE 428 AA; 48901 MW; 42E32733AD2B8F07 CRC64;
 CC -----
 CC Query Match 54.7%; Score 35; DB 1; Length 428;
 CC Best Local Similarity 54.5%; Pred. No. 53;
 CC Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 CC -----
 CC 1 NKG0QYTDQI 11
 CC |::|||
 CC Db 348 NONSRQYLDQI 358
 CC -----
 CC RESULT 18
 CC ID_HDA3 HUMAN STANDARD; PRT; 428 AA.
 CC AC O15379; O43268; Q9URF5; Q9URV0;
 CC DT 15-JUL-1998 (Rel. 36, Created)
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Histone deacetylase 3 (HD3) (RPD3-2).
 CC GN HDAC3.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OC NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A. (ISOFORM 1).
 CC RC TISSUE=Spleen, and T-cell;
 CC MEDLINE=98125547; PubMed=9464271;
 CC Dangond F., Hafler D.A., Tong J.K., Randall J., Kojima R., Utku N.,
 CC Gulians S.R.;
 CC "Differential display cloning of a novel human histone deacetylase
 CC (HDAC3) cDNA from PHA-activated immune cells";
 CC Biochem. Biophys. Res. Commun. 242:648-652(1998).
 CC RN [2]
 CC RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 CC RC TISSUE=Fibroblast;
 CC MEDLINE=98010646; PubMed=9346952;
 CC Yang W.-M., Yao Y.-L., Sun J.-M., Davie J.R., Seto E.;
 CC "Isolation and characterization of cDNAs corresponding to an
 CC additional member of the human histone deacetylase gene family";
 CC J. Biol. Chem. 272:28001-28007(1997).
 CC RN [3]
 CC RP SEQUENCE FROM N.A. (ISOFORM 1).
 CC RC MEDLINE=98196450; PubMed=9501169;
 CC RA Emiliani S., Fische W., van Lint C., Al-Abed Y., Verdin E.;
 CC "Characterization of a human RPD3 ortholog, HDAC3";
 CC Proc. Natl. Acad. Sci. U.S.A. 95:2795-2800(1998).
 CC RN [4]
 CC RP SEQUENCE FROM N.A. (ISOFORM 1).
 CC RC MEDLINE=99162406; PubMed=10051405;
 CC RA Mahlknecht U., Emiliani S., Naefeld V., Young S., Verdin E.;
 CC "Genomic organization and chromosomal localization of the human
 CC histone deacetylase 3 gene.";

RL Genomics 56:197-202(1999).
 RN [5]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Skin;
 RC MEDLINE=22386257; PubMed=12477932;
 RX Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hastef F.,
 RA Dichtchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Bawa S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 RA human and mouse cDNA sequences";
 RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RN SEQUENCE OF 95-353 AND 407-428 FROM N.A.
 RA Lynch E.D., Lee M.K., King M.-C.;
 RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RN INTERACTION WITH DAXX.
 RX MEDLINE=20136057; PubMed=10669754;
 RA Li H., Leo C., Zhu J., Wu X., O'Neill J., Park E.-J., Chen J.D.;
 RT "Sequestration and inhibition of Daxx-mediated transcriptional
 RT repression by PMW";
 RN Mol. Cell. Biol. 20:1794-1796(2000).
 RN [8]
 RN IDENTIFICATION IN A COMPLEX WITH NCOR1 AND NCOR2.
 RX PubMed=10860984;
 RA Wen Y.-D., Perissi V., Staszewski L.M., Yang W.-M., Krone A.,
 RA Glaes C.K., Rosenfeld M.G., Seto E.;
 RT "The histone deacetylase-3 complex contains nuclear receptor
 RT corepressors";
 RN Proc. Natl. Acad. Sci. U.S.A. 97:7202-7207(2000).
 RN [9]
 RN INTERACTION WITH HDAC7.
 RX PubMed=1146315;
 RA Fische W., Deguede F., Pillion M., Hendzel M.J., Voelter W.,
 RA Verdin E.;
 RT "Human HDAC7 histone deacetylase activity is associated with HDAC3 in
 RT vivo";
 RN J. Biol. Chem. 276:35826-35835(2001).
 RN [10]
 RN INTERACTION WITH HDAC10.
 RX MEDLINE=21851295; PubMed=11861901;
 RA Tong J.J., Liu J., Bertos N.R., Yang X.-J.;
 RT "Identification of HDAC10, a novel class II human histone deacetylase
 RT containing a leucine-rich domain";
 RN Nucleic Acids Res. 30:1114-1123(2002).
 RN [11]
 RN SUMOYLATION.
 RX PubMed=12032081;
 RA Kirsh O., Seeler J.-S., Pichler A., Gast A., Mueller S., Miska B.,
 RA Mathieu M., Harel-Bellan A., Kouraides T., Melchior F., Dejean A.;
 RT "The SUMO E3 ligase RanBP2 promotes modification of the HDACa
 RT deacetylase";
 RN EMBO J. 21:2682-2693(2002).
 CC -1- FUNCTION: Responsible for the deacetylation of lysine residues on
 CC the N-terminal part of the core histones (H2A, H2B, H3 and H4).
 CC Plays an important role in transcriptional regulation, cell cycle
 CC progression and developmental events. Histone deacetylases act via
 CC the formation of large multiprotein complexes. Probably

CC participates in the regulation of transcription through its
 CC binding to the zinc-finger transcription factor YY1; increases YY1
 CC repression activity. Required to repress transcription of the
 CC POU1F1 transcription factor.
 CC -1- SUBUNIT: Interacts with HDAC7. Forms a heterologous complex at
 CC with NCOR1 and NCOR2.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RPD3-2B;
 CC IsoId=O15379-1; Sequence=Displayed;
 CC Name=2; Synonyms=RPD3-2A;
 CC IsoId=O15379-2; Sequence=VSP_002079;
 CC -1- TISSUE SPECIFICITY: Widely expressed.
 CC -1- PTM: Sumoylated in vitro.
 CC -1- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
 CC FAMILY. HD SUBFAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U66914; AAC52038.1; -;
 CC EMBL: U75697; AAB8241.1; -;
 CC EMBL: U75696; AAB8240.1; -;
 CC EMBL: AF005482; AAB87752.1; -;
 CC EMBL: AF039703; AAC98927.1; -;
 CC EMBL: AF059650; AAC26509.1; -;
 CC EMBL: BC000614; AAH00614.1; -;
 CC EMBL: AF053138; AAC08351.1; -;
 CC EMBL: AF053137; AAC08351.1; JOINED.
 CC EMBL: AF053139; AAC08352.1; -;
 CC PIR: JC5834; JC5834; -;
 CC TRASNPA: T04110; -;
 CC Gene: HGNC:4854; HDAC3.
 CC MIM: 605166; -;
 CC GO: GO:0005634; C:nucleus; TAS.
 CC GO: GO:0008189; P:apoptosis inhibitor activity; TAS.
 CC GO: GO:0004407; P:histone deacetylase activity; TAS.
 CC GO: GO:0000074; P:regulation of cell cycle; TAS.
 CC InterPro: IPR000286; His_deacetylase; 1.
 CC Pfam: PF00850; Hist_deacetyl; 1.
 CC PRINTS: PR01270; HDASUPER.
 CC HydroLase; Nuclear protein; Chromatin regulator;
 CC Transcription regulation; Repressor; Ub1 conjugation;
 CC Alternative splicing.
 CC KEGG: K01101; HISTONE DEACETYLASE.
 CC FT DOMAIN 3 316
 CC FT ACT SITE 135 135
 CC FT VARSPPLIC 1 15
 CC FT BY SIMILARITY.
 CC FT MAAVAVFYDPDVG -> MIVKPYQASQHMCR (in
 CC isoform 2).
 CC FT CONFLICT 359 359 /Frid=VSP_002079.
 CC FT SEQUENCE 428 AA; 48847 MW; 94485C1EBDCFSAD0 CRC64;
 CC
 CC Query Match 54.7%; Score 35; DB 1; Length 428;
 CC Best Local Similarity 54.5%; Pred. No. 53;
 CC Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 1 NKGTQYTDQI 11
 CC ID : : : : :
 CC DB 348 NONSROYDOI 358
 CC
 CC RESULT 19
 CC ID LEUI_VIBCH STANDARD; PRT; 516 AA.
 CC AC Q9KPB3;
 CC DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 2-isopropylmalate synthase (EC 2.3.3.13) (Alpha-isopropylmalate
 DE synthase) (Alpha-IPM synthetase).
 GN LEUA OR VC2490.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Ueberback T., Fleischmann R.D., Mierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: Catalyzes the condensation of the acetyl group of
 CC acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
 CC 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O =
 CC 2-hydroxy-2-isopropylsuccinate + CoA.
 CC -1- PATHWAY: Leucine biosynthesis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate
 CC synthase family. LeuA 1 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE004318; AAF5632.1; -;
 CC PIR: F82070; F82070.
 CC TIGR: VC2490; -;
 CC HAMAP: MF_01025; -; 1.
 CC InterPro: IPR002034; AIPM/Hcit_synth.
 CC InterPro: IPR000891; HMGL-like.
 CC InterPro: IPR005671; LeuA_bact.
 CC Pfam: PF00682; HMGL-like; 1.
 CC TIGRPFAM: TIGR00973; LeuA_bact; 1.
 CC PROSITE: PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
 CC PROSITE: PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
 CC Leucine biosynthesis; transferase; Complete proteome.
 CC KEGG: K01101; TRANSFERASE.
 CC SEQUENCE 516 AA; 56163 MW; CDB7C04925887014 CRC64;
 CC
 CC Query Match 54.7%; Score 35; DB 1; Length 516;
 CC Best Local Similarity 54.5%; Pred. No. 65;
 CC Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 2 KGTQYTDQIE 12
 CC ID : : : : :
 CC DB 127 KHARGYTDVE 137
 CC
 CC RESULT 20
 CC ID YOXJ_BACSU STANDARD; PRT; 120 AA.
 CC AC P24809;
 CC DT 01-MAR-1992 (Rel. 21, Created)
 CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC GN Hypothetical protein yoxJ (ORF5).
 CC YOXJ OR YODF.

OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=9205258; PubMed=1683402;
 RA Foster S.J.;
 RT "Cloning, expression, sequence analysis and biochemical
 characterization of an autolytic amidase of Bacillus subtilis 168
 trpC2.";
 RL J. Gen. Microbiol. 137:1987-1998(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=95219086; PubMed=7704261;
 RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
 RT "Complete nucleotide sequence of a skin element excised by DNA
 rearrangement during sporulation in Bacillus subtilis.";
 RL Microbiology 141:323-327(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=97124195; PubMed=8969508;
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
 Kobayashi Y.;
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
 the Bacillus subtilis genome containing the skin element and many
 sporulation genes.";
 RL Microbiology 142:3103-3111(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azavedo V., Bertorello M.G., Bessieres P., Bolotin A., Borchert S.,
 Borries R., Bouteiller L., Brans A., Braum M., Brignelli S.C., Bron S.,
 Brouillet S., Brusch C.V., Caldwell I.F., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 Fritz G., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 Ghim S.Y., Guay B.J., Goffeau A., Golightly E.J., Grandi G.,
 Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Hbaut A.,
 Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Klaer-Bianchard M., Klein C.,
 Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 Kunita K., Lapides A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 Parro V., Pohl T.M., Portetelle D., Porrolik S., Prescott A.M.,
 Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 Rigter M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 Sato T., Scallan E., Schleich S., Schroeter R., Scofield F.,
 Setiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 Tanuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 Tostato V., Uchiyama S., Vandenol M., Vannier F., Vassarotti A.,
 Viari A., Wambutt R., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumbstein B., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.ebi-sib.ch/announce/>
 CC or send an email to license@ebi-sib.ch).

CC EMBL; M59232; AAA62678.1; -
 DR EMBL; D32216; BAA06962.1; -
 DR EMBL; D84432; BAA12426.1; -
 DR EMBL; Z99117; CAB14529.1; -
 DR PIR; D69968; D69968.
 DR Subtilist; Bg10507; YqXJ.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 120 AA; 14341 MW; AABD95706E20BD63 CRC64;
 Query Match 53.1%; Score 34; DB 1; Length 120;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 4 TOOYTDQIE 12
 Db 58 TSQYTDKIK 66
 RESULT 21
 UBIG_XYLFA STANDARD; PRT; 246 AA.
 AC OPRAMs;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3, 4-
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
 methyltransferase).
 GN UBIG OR XP2471.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_Taxid=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=945C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Renach P.C., Arruda P., Abreu F.A., Acencio M.,
 Alvarenga R., Alves L.M.C., Araya J.B., Bala G.S., Baptista C.S.,
 Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
 Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 Colaudo N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 Facchini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 Fraga J.S., Franca S.C., Franco M.C., Frome M., Furian L.R.,
 Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 Ho P.T., Holsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 Marques M.V., Martins E.A.L., Martins E.M.P., Matukuma A.Y.,
 Meuck C.F.M., Miraca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.R.S.,
 Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 de Oliveira M.C., de Oliveira R.C., Palieri D.A., Paris A.,
 de Oliveira R.B., Pereira G.A.G., Pereira H.A., Jr., Pasquero J.B.,
 Peixoto R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawaaki H.B.,
 da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
 Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 Zago M.A., Zatz M., Meidams J., Seubai J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -I- PATHWAY: Ubiquinone biosynthesis.
 CC -I- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AE004055; AA085269.1; -
DR PIR; H82553; H82553.
DR HAMAP; MF 00472; -; 1.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM_bind.
KW Ubiquitinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 246 AA; 27121 MW; 5212107D6363D3F CRC64;

Query Match 53.1%; Score 34; DB 1; Length 246;
Best Local Similarity 70.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

2 KSTQYTDQI 11
182 KSTHYKDFI 191

RESULT 22
ID UPS_METH STANDARD; PRT; 255 AA.

AC 026334;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) (UPP
DE synthetase) (Dl-trans-poly-cis-decaprenylcistransferase) (Undecaprenyl
DE di-phosphate synthase) (UDS).
GN MTH32.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_Taxid=187420;
RN [1]
RA SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldridge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RA strain: functional analysis and comparative genomics";
RT J. Bacteriol. 179:7135-7155(1997).

RL -1- FUNCTION: GENERATES UNDECAPRENYL PYROPHOSPHATE (UPP) FROM
CC ISOENENTYL PYROPHOSPHATE (IPP) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Di-trans-poly-cis-decaprenyl diphosphate +
CC isopentenyl diphosphate = diphosphate + di-trans-poly-cis-
CC undecaprenyl diphosphate.
CC -1- SIMILARITY: BELONGS TO THE UPP SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE000810; AAB84738.1; -
DR PIR; H69128; H69128.
DR InterPro; IPR001441; UPP_synth.
DR Pfam; PF01255; UPP_synthetase; 1.
DR Pfam; PD003461; UPP_synth; 1.

DR TIGRFAA; TIGR00055; upps; 1.
DR PROSITE; PS01066; UPP_SYNTHETASE; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 255 AA; 29900 MW; 370C6D8F85F8B8D CRC64;

Query Match 53.1%; Score 34; DB 1; Length 255;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

2 KSTQYTDQI 11
139 KSTQYSDRL 148

RESULT 23
ID PHC_STRAU STANDARD; PRT; 330 AA.

AC P09978;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospholipase C precursor (EC 3.1.4.3) (Beta-hemolysin) (Beta-toxin)
DE (Spningomycinase) (Smase).
GN HLB OR PLC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RA SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=89263748; PubMed=2726469;
RA Projan S.J., Kornblum J., Kreiswirth B., Moghazeh S.L., Eisner W.,
RA Novick R.P.;
RT "Nucleotide sequence: the beta-hemolysin gene of Staphylococcus
RT aureus";
RL Nucleic Acids Res. 17:3305-3305(1989).

RN [2]
RA SEQUENCE FROM N.A.
RC STRAIN=CN6708;
RX MEDLINE=91312132; PubMed=1830359;
RA Coleman D., Knights J., Russell R., Shanley D., Birkbeck T.H.,
RA Dougan G., Charles I.;
RT "Insertional inactivation of the Staphylococcus aureus beta-toxin by
RT bacteriophage phi 13 occurs by site- and orientation-specific
RT integration of the phi 13 genome";
RL Mol. Microbiol. 5:933-939(1991).

RN [3]
RA SEQUENCE FROM N.A.
RX MEDLINE=95028304; PubMed=7941863;
RA Katerov V.E., Golubkov V.I., Totolian A.A., Shalen K., Ensen L.,
RA Mikula I., Smola I.;
RT "The cloning and expression of the gene for Staphylococcus aureus
RT beta-hemolysin";
RL Zh. Mikrobiol. Epidemiol. Immunobiol. 3:28-33(1994).

RN [4]
RA SEQUENCE OF 35-53, AND CHARACTERIZATION.
RX STRAIN=RN4220;
RX MEDLINE=97072006; PubMed=8914839;
RA Dzielanowska K., Edwards V.M., Deringer J.R., Bohach G.A.,
RA Guerra D.J.;
RT "Comparison of the beta-toxins from Staphylococcus aureus and
RT Staphylococcus intermedius";
RL Arch. Biochem. Biophys. 335:102-108(1996).

CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE. BETA-HEMOLYSIN IS A
CC PHOSPHOLIPASE C WITH SPECIFIC ACTIVITY TOWARD SPHINGOMYELINS. HAS
CC A HIGH SPECIFICITY FOR SPHINGOMYELIN, HYDROLIZES
CC LYSPHOSPHATIDYLCHOLINE AT A MUCH LOWER RATE, BUT HAS NO ACTIVITY
CC TOWARDS PHOSPHATIDYLCHOLINE, PHOSPHATIDYLETHANOLAMINE, OR
CC PHOSPHATIDYLSERINE. THE PH OPTIMUM IS 6.5-7.5.
CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = 1,2-
CC diacylglycerol + choline phosphate.
CC -1- SUBUNIT: Monomer.

CC -1- SIMILARITY: BELONGS TO THE NEUTRAL SPHINGOMYELINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X13404; CAA31769.1; -
 DR EMBL: X61716; CAA43885.1; ALT_INIT.
 DR EMBL: S72497; AAB32218.1; ALT_INIT.
 DR PIR: S15324; S15324.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam: PF03372; Exo_endo_phos.1.
 KM Hemolysis; Hydrolyase; toxin; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 330 PHOSPHOLIPASE C.
 FT DISULFID 155 191 POTENTIAL.
 CC CONFLICT 51 51 Y -> T (IN REF. 4).
 CC SEQUENCE 330 AA, 37238 MW, 9F4148E23091020D CRC64;
 CC -----
 QY Query Match 53.1%; Score 34; DB 1; Length 330;
 Best Local Similarity 41.7%; Pred. No. 62;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 DB 1 NKGTOQYTDQI 12
 231 NKGTEPFDMK 242
 QY ||||| : : : : :
 DB ||||| : : : : :
 RESULT 24
 YOM5_CABEL STANDARD; PRT; 487 AA.
 AC P30651;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK643.5 in chromosome III.
 GN ZK643.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 STRAIN=Bristol N2;
 MEDLINE=92168156; PubMed=1538779;
 Sultoron J., Du Z., Thomas K., Wilson R., Hillier L., Staden R.,
 Halloran N., Green P., Thierry-Mieg J., Qiu L., Dear S., Coulson A.,
 Craxton M., Durbin R., Berks M., Metcalf M., Hawkins T.,
 RA "The C. elegans genome sequencing project: a beginning.";
 RA Nature 356:37-41(1992).
 RL [2]
 RN REVISIONS.
 RP Durbin R.;
 RA Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RL
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z1126; CAA77474.2; -
 DR WormBEP: ZK643.5; CE24732.
 KW Hypothetical protein.
 SQ SEQUENCE 487 AA, 55241 MW, BD6F6D02D7055013 CRC64;
 RT

QY Query Match 53.1%; Score 34; DB 1; Length 487;
 Best Local Similarity 60.0%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 DB 2 KGTQYTDQI 11
 38 KGTQESTDGL 47
 QY ||||| : : : : :
 DB ||||| : : : : :
 RESULT 25
 PURA_ARATH STANDARD; PRT; 490 AA.
 AC Q96529;
 ID PURA_ARATH
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenylosuccinate synthetase, chloroplast precursor (EC 6.3.4.4) (IMP--
 DE aspartate ligase) (AdSS) (AMPSase).
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 MEDLINE=96382483; PubMed=8790347;
 RA Fomne-Pfister R., Chemla P., Ward E., Girardet M., Kreuz K.E.,
 RA Honzatko R.B., Fromm H.J., Schaefer H.-P., Gruetter M.G.,
 RA Cowan-Jacob S.W.;
 RT "The mode of action and the structure of a herbicide in complex with
 RT its target: binding of activated adenylosuccinate synthetase.";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:9431-9436(1996).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=1130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,
 RA Patmanab B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,
 RA Delaeny M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Brothier P.,
 RA Wincker P., Cattolico L., Weisenbach J., Saurin W., Quetier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Bernes V.,
 RA Wurmbech E., Drzenek H., Erfle H., Jordan N., Bangert S., Nyakatura G.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordliek G.,
 RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,
 RA Cooke R., Landie M., Berger-Liauro C., Purnelle B., Masny D.,
 RA de Haan M., Maare A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argitlou A., Flores M., Lignori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Uterback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Maiz R., Wu D., Peterson J., Van Aken S.,
 RA Pail G., Miltcher J., Sellers P., Gill J.E., Feldlyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Frazer C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno M., Muraki A.,
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:820-822(2000).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 48-490.
 RX MEDLINE=20135989; PubMed=1065609;
 RA Prade L., Cowan-Jacob S.W., Chemla P., Potter S., Ward E.,
 RA Fomne-Pfister R.;
 RT "Structures of adenylosuccinate synthetase from Triticum aestivum and

RT Arabidopsis thaliana." ;
 RL J. Mol. Biol. 296:569-577(2000).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DE NOVO PATHWAY OF PURINE
 CC NUCLEOTIDE BIOSYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GMP + phosphate +
 CC adenylosuccinate.
 CC -1- PATHWAY: AMP biosynthesis; first committed step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U49389; AAB16828.1; -.
 CC EMBL; AL049660; CAB41194.1; -.
 CC PIR; T06759; T06759.
 CC PDB; 1D32; 24-MAR-00.
 DR Interpro; IPR001114; Asucc_synthetase.
 DR Pfam; PF00709; Adenylosucc_synth; 1.
 DR ProDom; PD001188; Asucc_synthetase; 1.
 DR TIGRPFam; TIGR00184; purA; 1.
 DR PROSITE; PS00513; ADENYLOSUCIN SYN 2; 1.
 DR PROSITE; PS01266; ADENYLOSUCIN SYN 1; 1.
 KW Purine biosynthesis; Ligase; GTP-binding; Chloroplast;
 KM Transic peptide; 3D-structure.
 FT TRANSIT 1 ?
 FT CHAIN ? 490
 FT NP BIND 77 83
 FT ACT SITE 206 206
 FT ACT SITE 213 213
 FT ACT SITE 213 213
 FT STRAND 63 64
 FT STRAND 68 73
 FT STRAND 81 88
 FT HELIX 89 91
 FT STRAND 94 97
 FT STRAND 102 103
 FT STRAND 106 109
 FT STRAND 115 118
 FT HELIX 123 126
 FT STRAND 128 129
 FT STRAND 131 134
 FT TURIN 136 137
 FT STRAND 139 140
 FT HELIX 142 153
 FT TURIN 154 156
 FT STRAND 160 162
 FT STRAND 163 166
 FT STRAND 170 171
 FT HELIX 174 186
 FT HELIX 199 208
 FT TURIN 209 209
 FT STRAND 213 213
 FT STRAND 214 218
 FT TURIN 220 222
 FT HELIX 223 227
 FT TURIN 239 240
 FT HELIX 245 262
 FT TURIN 263 265
 FT STRAND 266 267
 FT HELIX 269 278
 FT TURIN 279 280
 FT STRAND 283 290
 FT HELIX 291 293
 FT TURIN 295 297
 FT TURIN 300 301
 FT TURIN 310 310
 FT HELIX 311 315
 FT TURIN 316 318

FT TURIN 321 323
 FT STRAND 327 332
 FT STRAND 334 337
 FT TURIN 344 345
 FT STRAND 346 346
 FT HELIX 350 359
 FT TURIN 360 360
 FT STRAND 362 362
 FT TURIN 364 366
 FT STRAND 369 369
 FT STRAND 371 376
 FT HELIX 377 387
 FT TURIN 388 388
 FT STRAND 391 395
 FT HELIX 397 402
 FT STRAND 405 413
 FT STRAND 419 419
 FT HELIX 427 432
 FT STRAND 434 441
 FT TURIN 448 449
 FT TURIN 454 455
 FT HELIX 458 471
 FT STRAND 475 479
 FT STRAND 487 490
 FT SEQUENCE 490 AA; 52964 MW; B1E82BA386DF93CB CRC64;
 SQ
 Query Match 53.1%; Score 34; DB 1; Length 490;
 Best Local Similarity 54.5%; Pred. No. 95;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 KGTQGYDQIE 12
 DB 458 KAAQYVERIE 468
 RESULT 26
 Y538 RICPR STANDARD; PRT; 513 AA.
 ID Y538 RICPR
 AC Q9ZD12;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RP538.
 GN RP538.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OC Rickettsia; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sigheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria." ;
 RL Nature 396:133-140 (1998).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ235272; CAA14987.1; -.
 CC PIR; A71658; A71658.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 513 AA; 58705 MW; 68D24E1CFA71550D CRC64;
 SQ
 Query Match 53.1%; Score 34; DB 1; Length 513;

Best Local Similarity 50.0%; Pred. No. 99;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
3 GTCQYTDRE 12
431 GTCQYTDRE 440

RESULT 27

C3BB_BACTU ID C3BB_BACTU STANDARD; PRT; 652 AA.
AC 006117; 045717;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Pesticidal crystal protein cry3Bb (insecticidal delta-endotoxin
CryIIIBb) (crystalline entomocidal protoxin) (74 kDa crystal
protein).
CRY3BB OR CRYIIIB(B) OR CRYIIIB2.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1428;
[1]
SEQUENCE FROM N.A.
STRAIN=EG4961;
MEDLINE=93119147; PubMed=1476436;
Donovan W.P., Rupar M.J., Stanley A.C., Malvar T., Gawron-Burke M.C.,
Johnson T.B.;
"Characterization of two genes encoding Bacillus thuringiensis
insecticidal crystal proteins toxic to Coleoptera species.";
Appl. Environ. Microbiol. 58:3921-3927(1992).
[2]
SEQUENCE FROM N.A.
STRAIN=NR1-B-18655 / EG5144;
Donovan W.P., Rupar M.J., Stanley A.C.;
"Bacillus thuringiensis cryIIIC (b) protein toxic to coleopteran
insects.";
Patent number US5378625, 03-JAN-1995.
-1- FUNCTION: PROMOTES COLICIDIOSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPITHELIAL CELLS OF COLEOPTERA. HAS MODERATE LEVEL OF TOXICITY TO
SOUTHERN CORN ROOTWORM.
-1- SUBUNIT: Monomer.
-1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART OF
THE SPORE COAT.
-1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-
TERMINUS.
-1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M89794; AAA22334.1; -
EMBL; U31633; AAA74198.1; -
PIR; 139811; 139811.
DR PIR; 139811; 139811.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; Endotoxin C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
Toxin; Sporulation; 3D-structure.
FT VARIANT 21 21 O -> P (IN STRAIN EG5144).
FT VARIANT 97 97 I -> D (IN STRAIN EG5144).
FT VARIANT 289 289 I -> V (IN STRAIN EG5144).
FT VARIANT 352 352 F -> S (IN STRAIN EG5144).
FT VARIANT 417 417 VYL -> IYF (IN STRAIN EG5144).

FT VARIANT 451 451 S -> G (IN STRAIN EG5144).
FT VARIANT 590 590 L -> I (IN STRAIN EG5144).
FT VARIANT 600 600 K -> I (IN STRAIN EG5144).
FT VARIANT 624 624 K -> T (IN STRAIN EG5144).
SQ SEQUENCE 652 AA; 74385 MW; 6304832CDE8CC6 CRC64;

Query Match 53.1%; Score 34; DB 1; Length 652;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KTCQYTD 9
DB 235 KTCQYTD 242

RESULT 28

SAC7_YEAST ID SAC7_YEAST STANDARD; PRT; 654 AA.
AC P17121; 004163;
01-AUG-1990 (Rel. 15, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
DE GTPase-activating protein SAC7.
GN SAC7 OR YDR389W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
Dierich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
Carpenier J., Chen E., Cherry J.M., Chung E., Duncan M.,
Hunick-Smith S., Hyman R., Komp C., Laibkari D., Lew H., Lin D.,
Mosedale D., Nakamura K., Namath A., Oefner P., Oh C., Petel F.X.,
Roberts D., Schramm S., Schroeder M., Snogren T., Shroff N.,
Wintant A., Yelton M., Botstein D., Davis R.W.;
Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE OF 46-654 FROM N.A.
MEDLINE=90220616; PubMed=2183030;
Dunn T.M., Shortle D.;
"Null alleles of SAC7 suppress temperature-sensitive actin mutations
in Saccharomyces cerevisiae.";
Mol. Cell. Biol. 10:2308-2314(1990).
[3]
FUNCTION.
MEDLINE=97190189; PubMed=9038344;
Schmidt A., Bickel M., Beck T., Hall M.N.;
"The yeast phosphatidylinositol kinase homolog TOR2 activates RHO1 and
RHO2 via the exchange factor RIM2.";
Cell 88:531-542(1997).
-1- FUNCTION: GTPase activating protein for RHO1. Must be involved in
the normal assembly or function or both of actin. Plays an
essential role only at low temperatures.
-1- SIMILARITY: Contains 1 Rho-GAP domain.
-1- CAUTION: Ref.2 sequence differs from that shown due to a number of
framehifts.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U32274; AAB64831.1; -
EMBL; M32335; AAA5016.1; ALT_FRAME.
PIR; S69673; S69673.
DR S69673; S69673.
DR SGD; S0002797; SAC7.
DR InterPro; IPR00198; RHO GAP.
DR Pfam; PF00620; RHO GAP; 1.
DR SMART; SM00324; RHO GAP; 1.

DR PROSITE; PS50238; RHOGAP; 1.
KM GTPase activation; Cytokeleton.
FT DOMAIN 134 393 RHO-GAP.
SQ SEQUENCE 654 AA; 73629 MW; 50FD3387D13BBED CRC64;

Query Match 53.1%; Score 34; DB 1; Length 654;
Best Local Similarity 58.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KGTQOYTD 12
DB 415 NNGDKQKTPDIE 426

RESULT 29
ID C3BA_BACTO STANDARD; PRT; 659 AA.

AC P17969;
DT 01-NOV-1990 (Rel. 16, Created)
RT 01-NOV-1990 (Rel. 16, Last sequence update)
RL 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cry3Ba (insecticidal delta-endotoxin
CryIIIB(a)) (Crystalline entomocidal protoxin) (75 Kda crystal
protein).
GN CRY3BA OR CRYIIIB(A) OR CRYIIIB.
OS Bacillus thuringiensis (subsp. tolworthi).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1442;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=43F;
RX MEDLINE=90206811; PubMed=2320431;
RA Sick A., Gaertner F.H., Wong A.;
RT "Nucleotide sequence of a coleopteran-active toxin gene from a new
RT isolate of Bacillus thuringiensis subsp. tolworthi.";
RL Nucleic Acids Res. 18:1305-1305(1990).

CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPIHELIAL CELLS OF COLEOPTERA.

CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
OF THE SPORE COAT.

CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
N-TERMINUS.

CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X17123; CAA34983.1; -
DR EMBL; A07234; CAA00645.1; -
DR EMBL; S10228; S10228.

DR HSSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.

DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.

DR Pfam; PF03944; endotoxin_1.
DR Pfam; PF03945; endotoxin_C; 1.
DR Toxin; Sporulation.

SO SEQUENCE 659 AA; 75159 MW; 5A5B214FF84168CA CRC64;

Query Match 53.1%; Score 34; DB 1; Length 659;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 243 KLTQOYTD 250

RESULT 30

ID CLPB_SYNY3 STANDARD; PRT; 872 AA.

AC P74361;
DT 01-NOV-1997 (Rel. 35, Created)
RT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE CLPB protein.

GN CLPB OR SLR1641.

OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI_TaxId=1148;
RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Saito S., Kikuchi T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.,

RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).

CC -1- SUBUNIT: THOUGHT TO BE AN ATPASE SUBUNIT OF AN INTRACELLULAR
ATP-DEPENDENT PROTEASE (BY SIMILARITY)

CC -1- SIMILARITY: BELONGS TO THE CLPB/CLP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; D90914; BAA18456.1; -
DR PIR; S76197; S76197.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003599; AAA_ATPase_cent.

DR InterPro; IPR001270; Chaprinin_ClpA/B.

DR InterPro; IPR004176; Clp_N.

DR Pfam; PF00004; AAA; 2.

DR Pfam; PF02861; Clp_N; 2.

DR PRINTS; PRO0300; CLPPTREASEA.

DR SMART; SM00382; AAA; 2.

DR PROSITE; PS00870; CLPB_1; 1.

DR PROSITE; PS00871; CLPB_2; 1.

KW Chaperone; ATP-binding; Repeat; Complete proteome.

FT DOMAIN 163 411
FT DOMAIN 537 728
FT NP BIND 208 215
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

FT NP BIND 611 618
FT NP BIND 611 618
FT NP BIND 611 618

RN [10]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.t.alzawai; STRAIN=IC1;
 RX MEDLINE=89083518; PubMed=3205732;
 RA Halder M.Z., Ellar D.J.;
 RT "Nucleotide sequence of a *Bacillus thuringiensis* aizawai IC1
 entomocidal crystal protein gene";
 RL Nucleic Acids Res. 16:10927-10927(1988).
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPIHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORE/STATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- BIOCHEMISTRY: Introduced by genetic manipulation and expressed in
 CC insect-resistant maize by Monsanto, Northrup King and Ciba Geigy.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; M15271; AAA22561.1; -
 CC EMBL; M13898; AAA22330.1; -
 CC EMBL; M37263; AAA22420.1; -
 CC EMBL; M12661; AAA22613.1; -
 CC EMBL; AF059670; AAC64003.1; -
 CC EMBL; D00117; BAA00071.1; -
 CC EMBL; X04698; CAA28405.1; -
 CC EMBL; X54939; CAA38701.1; -
 CC EMBL; A09398; CAA00840.1; -
 CC EMBL; A03793; CAA00303.1; -
 CC EMBL; M1463; AAA22551.1; -
 CC EMBL; X13233; CAA31620.1; -
 CC PIR; A26513; A26513.
 CC PIR; A29838; A29838.
 CC PIR; A90025; ID0002.
 CC PIR; I39838; I39838.
 CC PIR; S02134; S02134.
 CC HSSP; P02965; IC1Y.
 CC InterPro: IPR001178; Endotoxin.
 CC InterPro: IPR005638; endotoxin_C.
 CC InterPro: IPR005639; endotoxin_N.
 CC Pfam: PF03944; endotoxin_C; 1.
 CC Pfam: PF03945; endotoxin_N; 1.
 KM Toxin; Sporulation; Plasmid; Genetically modified food.
 FT CONFLICT 207 283
 FT CONFLICT 382 406
 FT CONFLICT 410 410
 FT CONFLICT 430 432
 FT CONFLICT 437 437
 FT CONFLICT 447 447
 FT CONFLICT 450 450
 FT CONFLICT 452 458
 FT CONFLICT 461 461
 FT CONFLICT 463 465
 FT CONFLICT 479 486
 FT CONFLICT 492 492
 FT CONFLICT 501 505
 FT CONFLICT 537 537
 FT CONFLICT 542 542
 FT CONFLICT 545 545
 FT CONFLICT 568 568
 FT CONFLICT 569 569
 FT CONFLICT 665 665
 K -> E (IN REF. 3).
 L -> P (IN REF. 4).
 LSH -> CLAY (IN REF. 4).
 R -> Y (IN REF. 4).
 I -> V (IN REF. 4).
 A -> P (IN REF. 10).
 MFSWIR -> NDSWYTC (IN REF. 4).
 E -> N (IN REF. 3).
 NNI -> GOV (IN REF. 4).
 STUUGSGT -> LOSWLMN (IN REF. 4).
 P -> L (IN REF. 4).
 RRTSP -> ELTF (IN REF. 4).
 F -> L (IN REF. 10).
 D -> H (IN REF. 3).
 P -> I (IN REF. 10).
 T -> I (IN REF. 10).
 TV -> HL (IN REF. 3).
 K -> E (IN REF. 3).

FT CONFLICT 675 676 KR -> NG (IN REF. 3).
 FT CONFLICT 703 703 S -> N (IN REF. 4).
 FT CONFLICT 712 712 D -> H (IN REF. 4).
 FT CONFLICT 731 731 P -> L (IN REF. 7 AND 10).
 FT CONFLICT 785 785 P -> R (IN REF. 7 AND 10).
 FT CONFLICT 836 836 N -> I (IN REF. 3).
 FT CONFLICT 978 978 H -> Q (IN REF. 9).
 FT CONFLICT 1016 1016 E -> T (IN REF. 3).
 FT CONFLICT 1036 1036 C -> F (IN REF. 9).
 FT CONFLICT 1060 1060 E -> G (IN REF. 3).
 SQ SEQUENCE 1155 AA; 130623 MW; 43461A64C7AC7CAF CRC64;
 Query Match 53.1%; Score 34; DB 1; Length 1155;
 Best Local Similarity 50.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 NKGTOOYDQTE 12
 DB 592 NSGNEVYIDRIE 603
 RESULT 33
 ID C1A_BACTK STANDARD; PRT; 1176 AA.
 AC P02965; P09664; P09665; P16478; Q9RED5;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pesticidal crystal protein cryIIa (insecticidal delta-endotoxin
 DE CryI(a)) (Crystalline entomocidal protoxin) (113 kDa crystal protein).
 GN CRYIIA OR CRYIA(A) OR CRYIA OR CRY-1-1 OR CRYA OR ICP OR CRYBNS3-1.
 OS *Bacillus thuringiensis* (subsp. kurstaki),
 OS *Bacillus thuringiensis* (subsp. aizawai),
 OS *Bacillus thuringiensis* (subsp. entomocidus), and
 OS *Bacillus thuringiensis* (subsp. sotto).
 OG Plasmid 68 kb.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=29339, 1433, 1436, 29340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.t.kurstaki; STRAIN=HD-1-Dippei;
 RX MEDLINE=85207613; PubMed=2581950;
 RA Schnepf H.E., Wong H.C., Whiteley H.R.;
 RT "The amino acid sequence of a crystal protein from *Bacillus*
 RT *thuringiensis* deduced from the DNA base sequence.";
 RL J. Biol. Chem. 260:6264-6272(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.t.kurstaki; STRAIN=BNS3;
 RA Tounsi S., J'Mai A., Zouari N., Jaoua S.;
 RT "Cloning and nucleotide sequence of a novel cryIIa-type gene from
 RT *Bacillus thuringiensis* subsp. kurstaki.";
 RL Biotechnol. Lett. 21:771-775(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.t.alzawai; STRAIN=IPV7;
 RA Simizu M., Oshie K., Nakamura K., Takada Y., Oeda K.;
 RT "Cloning and expression in *Escherichia coli* of the 135-kDa
 RT insecticidal protein gene from *Bacillus thuringiensis* subsp. aizawai
 RT IPV7.";
 RL Agric. Biol. Chem. 52:1565-1573(1988).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=B.t.entomocidus;
 RX MEDLINE=89098405; PubMed=2911478;
 RA Masson V., Marcotte P., Prefontaine G., Brousseau R.;
 RT "Nucleotide sequence of a gene cloned from *Bacillus thuringiensis*
 RT subsp. entomocidus coding for an insecticidal protein toxic for
 RT *Bombyx mori*.";
 RL Nucleic Acids Res. 17:446-446(1989).
 RN [5]
 RP SEQUENCE OF 1-934 FROM N.A.
 RC SPECIES=B.t.sotto;

RX MEDLINE=85232070; PubMed=2989108; Shihano Y., Yamagata A., Nakamura N., Iizuka T., Sugisaki H., Takamami M.;
RA "Nucleotide sequence coding for the insecticidal fragment of the
RT Bacillus thuringiensis crystal protein.",
RL Gene 34:243-251(1985).
RN
RP [6]
RQ SEQUENCE OF 1-333 FROM N.A.
RC SPECIES=B.t.kurstaki; STRAIN=HD-1-Dippel;
RX MEDLINE=83109004; PubMed=6296116;
RA Wong H.C., Schenep H.E., Whiteley H.R.;
RT "Transcriptional and translational start sites for the Bacillus
RL thuringiensis crystal protein gene.",
RN J. Biol. Chem. 258:1960-1967(1983).
RQ
RN [7]
RX X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 33-609.
RC SPECIES=B.t.kurstaki; STRAIN=HD-1-Dippel;
RX MEDLINE=96095803; PubMed=7490762;
RA Grochulski P., Masson L., Borissova S., Pusztai-Carey M.,
Schwartz J.L., Brousseau R., Cygler M.;
RT "Bacillus thuringiensis CryIA(a) insecticidal toxin: crystal
RL structure and channel formation.",
RN J. Mol. Biol. 254:447-464(1995).
RQ
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPIHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.
CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01554; -; NOT ANNOTATED_CDS.
DR EMBL; Y09663; CAA70856.1; -;
DR EMBL; M1250; AAA22353.1; -;
DR EMBL; D00348; BAA00257.1; -;
DR EMBL; M10917; AAA22552.1; -;
DR EMBL; X13355; CAA1886.1; -;
DR PIR; A22617; A22617.
DR PIR; J02041; J02041.
DR PIR; S02215; S02215.
DR PDB; 1C1Y; 27-JAN-97.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation; Plasmid; 3D-structure.
FT VARIANT 77 P -> L (IN STRAINS BNS3, IPL7 AND SOTTO).
FT VARIANT 148 L -> F (IN STRAINS IPL7 AND SOTTO).
FT VARIANT 302 S -> R (IN STRAIN SOTTO).
FT VARIANT 918 Q -> R (IN STRAIN SOTTO).
FT CONFLICT 1009 V -> L (IN REF. 1).
FT HELIX 35 48
FT HELIX 54 63
FT TURN 64 65
FT HELIX 70 84
FT TURN 85 85
FT HELIX 90 119
FT TURN 121 122
FT HELIX 124 144
FT HELIX 145 148
FT TURN 151 152
FT HELIX 154 178

FT HELIX 180 182
FT TURN 183 183
FT HELIX 186 218
FT HELIX 223 239
FT TURN 240 240
FT HELIX 241 244
FT TURN 245 246
FT HELIX 247 250
FT TURN 252 254
FT STRAND 259 260
FT STRAND 266 269
FT HELIX 271 274
FT TURN 281 282
FT HELIX 284 289
FT TURN 290 290
FT STRAND 298 310
FT TURN 311 312
FT STRAND 313 325
FT HELIX 326 328
FT STRAND 333 334
FT STRAND 338 339
FT STRAND 345 351
FT TURN 354 355
FT STRAND 357 367
FT TURN 375 376
FT STRAND 380 390
FT STRAND 400 402
FT STRAND 408 409
FT HELIX 410 412
FT TURN 413 413
FT HELIX 419 420
FT TURN 423 426
FT STRAND 429 434
FT STRAND 437 438
FT TURN 441 442
FT STRAND 446 449
FT STRAND 452 456
FT TURN 457 458
FT STRAND 464 465
FT STRAND 470 474
FT HELIX 475 477
FT STRAND 480 481
FT TURN 483 484
FT STRAND 486 488
FT STRAND 498 501
FT STRAND 505 514
FT TURN 518 519
FT STRAND 522 530
FT STRAND 534 540
FT TURN 541 542
FT STRAND 543 550
FT TURN 556 557
FT HELIX 562 564
FT STRAND 566 569
FT STRAND 574 574
FT STRAND 580 588
FT TURN 592 593
FT STRAND 596 605
FT TURN 606 607
SQ SEQUENCE 1176 AA; 133119 NM; E2BE15AF12E5DD85 CRC64;
Query Match 53.1%; Score 34; DB 1; Length 1176;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQIE 12
DB 591 NSGNEVYIDRIE 602
RESULT 34
C1AG_BACTU STANDARD; PRT; 1176 AA.
ID C1AG_BACTU

AC G95515;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Peptidicidal crystal protein cryIaG (Insecticidal delta-endotoxin
 DE cryIa(g)) (crystalline entomocidal protoxin) (134 kDa crystal protein).
 GN CRYIaG OR CRYIa(G) OR CRYI.
 OS Bacillus thuringiensis.
 OC Bacterioidetes; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mustafa S.A.;
 RT "Cloning, nucleotide sequence and expression of a new gene encoding
 RT lepidopteran-specific toxin";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORELIATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: A081248; AAD46137.1; -
 DR HSSP: P02965; 1C1Y.
 DR InterPro: IPR001178; Endotoxin.
 DR InterPro: IPR005638; endotoxin_C.
 DR InterPro: IPR005639; endotoxin_N.
 DR Pfam: PF00555; endotoxin_1.
 DR Pfam: PF03944; endotoxin_C; 1.
 DR Pfam: PF03945; endotoxin_N; 1.
 DR Toxin; Sporulation; Plasmid.
 KW SEQUENCE 1176 AA; 133392 MW; 285929160A2612BB CRC64;
 SQ
 Query Match 53.1%; Score 34; DB 1; Length 1176;
 Best Local Similarity 50.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 1 NKGTQOYTDQIE 12
 DB 591 NSGNEYVDRIE 602
 RESULT 35
 C1AE_BACTL STANDARD; PRT; 1181 AA.
 ID C1AE_BACTL
 AC 003748;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Peptidicidal crystal protein cryIaE (Insecticidal delta-endotoxin
 DE cryIa(e)) (crystalline entomocidal protoxin) (134 kDa crystal protein).
 GN CRYIaE OR CRYIa(E) OR ENDI.
 OS Bacillus thuringiensis (subsp. alesti).
 OC Bacterioidetes; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1440;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee C.S.; Aronson A.I.;
 RT "Cloning and analysis of delta-endotoxin genes from Bacillus
 RT thuringiensis subsp. alesti.";

RL J. Bacteriol. 173:6635-6638(1991).
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORELIATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M65252; AAA22410.1; -
 DR PIR: A41052; A41052.
 DR HSSP: P02965; 1C1Y.
 DR InterPro: IPR001178; Endotoxin.
 DR InterPro: IPR005638; endotoxin_C.
 DR InterPro: IPR005639; endotoxin_N.
 DR Pfam: PF00555; endotoxin_1.
 DR Pfam: PF03944; endotoxin_C; 1.
 DR Pfam: PF03945; endotoxin_N; 1.
 DR Toxin; Sporulation.
 KW SEQUENCE 1181 AA; 133737 MW; FE86D2138C37D0FD CRC64;
 SQ
 Query Match 53.1%; Score 34; DB 1; Length 1181;
 Best Local Similarity 50.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 1 NKGTQOYTDQIE 12
 DB 592 NSGNEYVDRIE 603
 RESULT 36
 UB2_KULUA STANDARD; PRT; 1221 AA.
 ID UB2_KULUA
 AC 042726;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 2 (EC 3.1.2.15) (Ubiquitin
 DE thiolesterase 2) (Ubiquitin-specific processing protease 2)
 DE (Deubiquitinating enzyme 2).
 GN UB2.
 OS Kluveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JBD100;
 RA MEDLINE=20407965; PubMed=10953877;
 RA Winkler A.A.; Korfstamje R.; Zonneveld B.J.M.; Hooykaas P.J.J.;
 RA Steensma H.Y.;
 RT "Isolation and characterization of KUBP2, a ubiquitin hydrolase gene
 RT of Kluveromyces lactis that can suppress a ts-mutation in CBF2, a
 RT gene encoding a centromeric protein of Saccharomyces cerevisiae.";
 RL Curr. Genet. 38:17-22(2000).
 CC -1- FUNCTION: HAS AN ATP-INDEPENDENT ISOPEPTIDASE ACTIVITY, CLEAVING
 CC AT THE CARBOXYL TERMINUS OF THE UBIQUITIN MOIETY IN NATURAL OR
 CC ENGINEERED LINEAR FUSION PROTEINS, IRRESPECTIVE OF THEIR SIZE OR
 CC THE PRESENCE OF AN AMINO-TERMINAL EXTENSION TO UBIQUITIN (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -1- SIMILARITY: Belongs to peptidase family C19.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF022776; AAB94074.1; -
DR PIR; T30529; T30529.
DR MEROPS; C19.003; -
DR InterPro; IPR001394; UCH-2.
DR Pfam; PF00443; UCH-1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
DR Ubl conjugation pathway; Hydrolyase; Thiol protease; Multigene family.
FT ACT_SITE 707 BY SIMILARITY.
FT ACT_SITE 1146 BY SIMILARITY.
FT ACT_SITE 1154 BY SIMILARITY.
SEQUENCE 1221 AA; 141474 MW; 2001D3A67263CA62 CRC64;

Query Match 53.1%; Score 34; DB 1; Length 1221;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTQGYTDQI 11
|||:|:|:
Db 857 GTQGYSVQV 865

RESULT 37
MYH4 RABIT STANDARD; PRT; 1938 AA.

AC Q28641; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DR Myosin heavy chain, skeletal muscle, juvenile.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Skeletal muscle;
RA Maeda K., Hosinova E., Roesch-Kleinhauf A., Schuster H., Gaepertik J.,
Wittinghofer A.;
RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
skeletal muscle and a novel cosynthesis of S-1 fragment with the
essential and regulatory light chains."
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAINS: THE ROD-LIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEMORMYOSIN (LMH) AND 1 HEAVY MEMORMYOSIN (HMH). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -1- SIMILARITY: Contains 1 IQ domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; U32574; AAA74199.1; -
DR PIR; A59293; A59293.
DR HSSB; P13538; 2MTS.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; myosin_N.
DR InterPro; IPR002928; myosin_tail.
DR Pfam; PF00612; IQ_2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SMO0035; IQ_1.
DR SMART; SMO0242; MISC; 1.

DR PROSITE; PS50096; IQ_1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
FT DOMAIN 784 813 IQ.
FT DOMAIN 842 1938 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 760 774 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
FT MOD_RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 756 756 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD_RES 698 698 ALKYLATION (SH-1) (BY SIMILARITY).
FT MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
SQ SEQUENCE 1938 AA; 223064 MW; D8A8A2BC5B182626 CRC64;

Query Match 53.1%; Score 34; DB 1; Length 1938;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
:::|:|:|:
Db 1305 SRGQAFQIQIE 1316

RESULT 38
MYH1 HUMAN STANDARD; PRT; 1939 AA.

AC P12882; Q9Y622; 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DR Myosin heavy chain, skeletal muscle, adult 1 (Myosin heavy chain
IIX/d) (MYHC-IIX/d).
GN MYH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
heavy chain family: implications for functional diversity.";
RT J. Mol. Biol. 290:61-75(1999).
RN [2]

CC SEQUENCE OF 1064-1939 FROM N.A.
CC MEDLINE=86176778; PubMed=2421254;
CC Saez L., Leinwand L.A.;
CC "Characterization of diverse forms of myosin heavy chain expressed in
CC adult human skeletal muscle.";
CC Nucleic Acids Res. 14:2951-2969(1986).

CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -1- SIMILARITY: Contains 1 IQ domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF111785; AAD29951.1; -.
 CC EMBL: X03740; CAA27380.1; -.
 CC PIR: A23767; A23767.
 CC HSSP: P13538; 2MTS.
 CC DR HSP: HSPC7567; MYTH1.
 CC DR MIM: 160730; -.
 CC DR InterPro: IPR000048; IQ_region.
 CC DR InterPro: IPR001609; Myosin_head.
 CC DR InterPro: IPR004009; Myosin_N.
 CC DR InterPro: IPR002928; Myosin_tail.
 CC Pfam: PF00612; IQ; 2.
 CC Pfam: PF00663; myosin_head; 1.
 CC Pfam: PF02736; Myosin_N; 1.
 CC Pfam: PF01576; Myosin_tail; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART: SM00015; IQ; 1.
 CC SMART: SM00242; MYSC; 1.
 CC PROSITE: PS50096; IQ; 1.
 CC Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 CC Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 CC Multigene family.
 CC -----
 CC DOMAIN 1 784
 CC 1 785 814
 CC DOMAIN 843 1939
 CC NP_BIND 179 186
 CC DOMAIN 659 681
 CC DOMAIN 761 775
 CC MOD_RES 130 130
 CC MOD_RES 699 699
 CC MOD_RES 709 709
 CC FT CONFLICT 1131 1131
 CC FT CONFLICT 1139 1139
 CC FT CONFLICT 1158 1158
 CC FT CONFLICT 1163 1163
 CC FT CONFLICT 1286 1289
 CC FT CONFLICT 1302 1303
 CC FT CONFLICT 1451 1451
 CC FT CONFLICT 1470 1470
 CC FT CONFLICT 1474 1474
 CC FT CONFLICT 1569 1569
 CC FT CONFLICT 1598 1598
 CC FT CONFLICT 1606 1606
 CC FT CONFLICT 1643 1643
 CC FT CONFLICT 1648 1648
 CC FT CONFLICT 1750 1750
 CC FT CONFLICT 1822 1822
 CC FT CONFLICT 1845 1845
 CC -----
 CC MYOSIN HEAD-LIKE.
 CC IQ.
 CC COILED COIL (POTENTIAL).
 CC ATP (POTENTIAL).
 CC ACTIN-BINDING (BY SIMILARITY).
 CC ACTIN-BINDING (BY SIMILARITY).
 CC METHYLATION (SH-1) (POTENTIAL).
 CC ALKYLATION (SH-2) (POTENTIAL).
 CC ALKYLATION (SH-2) (POTENTIAL).
 CC A -> T (IN REF. 2).
 CC Q -> L (IN REF. 2).
 CC G -> V (IN REF. 2).
 CC A -> T (IN REF. 2).
 CC TEGS -> ONOV (IN REF. 2).
 CC VS -> ET (IN REF. 2).
 CC R -> T (IN REF. 2).
 CC E -> V (IN REF. 2).
 CC AS -> SF (IN REF. 2).
 CC L -> V (IN REF. 2).
 CC V -> N (IN REF. 2).
 CC D -> E (IN REF. 2).
 CC A -> Q (IN REF. 2).
 CC R -> Q (IN REF. 2).
 CC Q -> K (IN REF. 2).
 CC R -> K (IN REF. 2).
 CC R -> H (IN REF. 2).

SQ SEQUENCE 1939 AA; 223114 MW; 39ADB26AB79DFA53 CRC64;
 Query Match 53.1%; Score 34; DB 1; Length 1939;
 Best Local Similarity 50.0%; Pred. No. 4.1e-02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MKGTQYTDQIE 12
 DB 1306 SRGKAFTQIE 1317
 RESULT 39
 MYH4 HUMAN
 ID MYH4 HUMAN STANDARD; PRT; 1939 AA.
 AC Q9Y623;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain IIB)
 DE (MyHC-IIb).
 GN MYH4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=skeletal muscle;
 RX MEDLINE=99318869; PubMed=10388558;
 RA Weiss A., Schiaffino S., Lainwand L.A.;
 RT "Comparative sequence analysis of the complete human sarcomeric myosin
 RT heavy chain family: implications for functional diversity."
 RL J. Mol. Biol. 290:61-75(1999).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -1- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF111783; AAD29949.1; -.
 CC HSSP: P13538; 2MTS.
 CC DR HSP: HSPC7574; MYTH4.
 CC DR MIM: 160742; -.
 CC DR InterPro: IPR000048; IQ_region.
 CC DR InterPro: IPR001609; myosin_head.
 CC DR InterPro: IPR004009; Myosin_N.
 CC DR InterPro: IPR002928; Myosin_tail.
 CC Pfam: PF00612; IQ; 2.
 CC Pfam: PF00663; myosin_head; 1.
 CC Pfam: PF02736; Myosin_N; 1.
 CC Pfam: PF01576; Myosin_tail; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART: SM00015; IQ; 1.

```

DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
FT DOMAIN 785 814 IQ.
FT DOMAIN 843 1939 COILED COIL (POTENTIAL).
FT NP BIND 179 186 ATP (POTENTIAL).
SQ SEQUENCE 1939 AA; 223012 MW; 40B1AD1D777A47DE CRC64;

Query Match 53.1%; Score 34; DB 1; Length 1939;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
DB 1306 SRGQAFQDQIE 1317

ULT 40
H3 CHICK
ID MYH3 CHICK STANDARD; PRT; 1940 AA.
AC P02565;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87194881; PubMed=3571266;
RA Molina M.I., Kropp K.E., Gulick J., Robbins J.;
RT "The sequence of an embryonic myosin heavy chain gene and isolation
RT of its corresponding cDNA."
RL J. Biol. Chem. 262:6478-6486 (1987).
RN [2]
RP SEQUENCE OF 1502-1940 FROM N.A.
RX MEDLINE=83161144; PubMed=6833296;
RA Kavinsky C.J., Umeda P.K., Sinha A.M., Elzinga M., Tong S.W., Zak R.,
RA Jakovcic S., Rabinowitz M.,
RT "Cloned mRNA sequences for two types of embryonic myosin heavy chains
RT from chick skeletal muscle. I. DNA and derived amino acid sequence of
RT light meromyosin."
J. Biol. Chem. 258:5196-5205 (1983).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -1- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; V00430; CA23712.1; -
CC EMBL; J02714; AAA48972.1; -
CC PIR; A29320; A29320.
CC HSSP; P13538; 2MYS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR004009; myosin_N.
CC InterPro; IPR002828; myosin_tail.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF02736; Myosin N; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD00035; myosin_head; 1.
CC DR SMART; SM00015; IQ; 1.
CC DR SMART; SM00242; MYSC; 1.
CC DR PROSITE; PS50096; IQ; 1.
CC KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
CC KW Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 844 1940 COILED COIL (POTENTIAL).
FT NP BIND 179 186 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 762 776 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1).
FT MOD_RES 710 710 ALKYLATION (SH-2).
FT VARIANT 379 379 G -> D.
FT CONFLICT 1547 1547 T -> A (IN REF. 2).
FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
SQ SEQUENCE 1940 AA; 222816 MW; C34833D75B04DF2 CRC64;

Query Match 53.1%; Score 34; DB 1; Length 1940;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
DB 1307 SRGQAFQDQIE 1318

```

Search completed: August 20, 2003, 09:30:41
Job time : 8.92783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 34.8866 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-4
Perfect score: 64
Sequence: 1 NKGTQYTDQIE 12

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaeal: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	138	12	Q85155 human parvo
2	64	100.0	141	12	Q85171 human parvo
3	64	100.0	141	12	Q85166 human parvo
4	64	100.0	141	12	Q85146 human parvo
5	64	100.0	142	12	Q85168 human parvo
6	64	100.0	142	12	Q85173 human parvo
7	64	100.0	144	12	Q85138 human parvo
8	64	100.0	145	12	Q85181 human parvo
9	64	100.0	145	12	Q85161 human parvo
10	64	100.0	146	12	Q85158 human parvo
11	64	100.0	147	12	Q85142 human parvo
12	64	100.0	148	12	Q85150 human parvo
13	64	100.0	151	12	Q85131 human parvo
14	64	100.0	151	12	Q85177 human parvo
15	64	100.0	153	12	Q85123 human parvo
16	64	100.0	157	12	Q85196 human parvo

17	64	100.0	162	12	Q85135 human parvo
18	64	100.0	546	12	Q913X0 human parvo
19	64	100.0	546	12	Q913W7 human parvo
20	64	100.0	554	12	Q9PZS9 human parvo
21	64	100.0	554	12	Q90201 human parvo
22	64	100.0	554	12	Q65790 human parvo
23	64	100.0	554	12	Q912B7 human eryth
24	64	100.0	554	12	Q9JGP7 human parvo
25	64	100.0	554	12	Q9WKL9 human parvo
26	64	100.0	554	12	Q8JN54 human parvo
27	64	100.0	760	12	Q9PZT8 human parvo
28	64	100.0	761	12	Q9PZU0 human parvo
29	64	100.0	765	12	Q9PZT6 human parvo
30	64	100.0	769	12	Q9PZT4 human parvo
31	64	100.0	773	12	Q913X1 human parvo
32	64	100.0	773	12	Q913W8 human parvo
33	64	100.0	781	12	Q8JN56 human parvo
34	64	100.0	781	12	Q65789 human parvo
35	64	100.0	781	12	P90223 human parvo
36	64	100.0	781	12	Q85191 human parvo
37	64	100.0	781	12	P90221 human parvo
38	64	100.0	781	12	P89318 human parvo
39	64	100.0	781	12	Q8JN53 erythrocyte
40	64	100.0	781	12	P90224 human parvo
41	64	100.0	781	12	Q9PZT0 human parvo
42	64	100.0	781	12	P89316 human parvo
43	64	100.0	781	12	P89320 human parvo
44	64	100.0	781	12	Q85117 human parvo
45	64	100.0	781	12	P89321 human parvo

ALIGNMENTS

RESULT 1
ID Q85155 PRELIMINARY; PRT; 138 AA.
AC Q85155;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VPI.
OS Human parvovirus B19.
OC Viruses: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Pöblitzki A., Giegler A., Cassinotti P., Siegl G., Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates."
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70564; CAA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899FB879A3B68D CRC64;
FT
Query Match 100.0%; Score 64; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTQYTDQIE 12
DB 107 NKGTQYTDQIE 118
RESULT 2
ID Q85171 PRELIMINARY; PRT; 141 AA.
AC Q85171;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 270580; CAA94493.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT SEQUENCE 141 AA; 15770 MW; C9E92572A78C6C27 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 0.00074;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDOIE 12
 Db 109 NKGTOOYTDOIE 120
 RESULT 3
 ID 085166 PRELIMINARY; PRT; 141 AA.
 AC 085166;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 270575; CAA94488.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT SEQUENCE 141 AA; 15785 MW; 4115D3D915751757 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 0.00074;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDOIE 12
 Db 109 NKGTOOYTDOIE 120
 RESULT 4
 ID 085146 PRELIMINARY; PRT; 141 AA.
 AC 085146;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 270555; CAA94467.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 141
 FT SEQUENCE 141 AA; 15712 MW; 32F6B23E7E6B8583 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 141;
 Best Local Similarity 100.0%; Pred. No. 0.00074;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDOIE 12
 Db 106 NKGTOOYTDOIE 117
 RESULT 5
 ID 085168 PRELIMINARY; PRT; 142 AA.
 AC 085168;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 270577; CAA94490.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 142
 FT SEQUENCE 142 AA; 15821 MW; 2B6B4D9A6784F8C4 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 0.00075;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDOIE 12
 Db 105 NKGTOOYTDOIE 116
 RESULT 6
 ID 085173 PRELIMINARY; PRT; 142 AA.
 AC 085173;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE Viral protein 1 (Fragment).
 GN VPI.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 ON NCBI_TaxId=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,

RA Wolf H., Modrow S.,
RT "XXXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70567; CA94480.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 142
SQ SEQUENCE 142 AA; 15820 MW; 3A95E07C0BC0434 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 105 NKGTQOYTDQIE 116

RESULT 7
AC 085138 PRELIMINARY; PRT; 144 AA.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
RT "XXXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70547; CA94459.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 144
SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AE CRC64;

Query Match 100.0%; Score 64; DB 12; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQIE 12
107 NKGTQOYTDQIE 118

RESULT 8
AC 085181 PRELIMINARY; PRT; 145 AA.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
RT "XXXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70590; CA94503.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16210 MW; 6B45A1EB9B23C4 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 109 NKGTQOYTDQIE 120

RESULT 9
AC 085161 PRELIMINARY; PRT; 145 AA.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
RT "XXXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70570; CA94483.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQIE 12
105 NKGTQOYTDQIE 116

RESULT 10
AC 085158 PRELIMINARY; PRT; 146 AA.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Pobioltzki A., Gigler A., Cassinotti P., Siegl G.,
RT "XXXXSequence variability among different parvovirus B19 isolates."
RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70567; CA94480.1; -
DR InterPro: IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 146
SQ SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6E4D9A CRC64;

Query Match 100.0%; Score 64; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 105 NKGTQOYTDQIE 116

RESULT 11

ID 085142 PRELIMINARY; PRT; 147 AA.

AC 085142; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DR Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,

Wolf H., Modrow S.; "XXXSequence variability among different parvovirus B19 isolates.";

RT Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RL EMBL; Z70551; CAA94463.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 147 147

SQ SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 147;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 108 NKGTQOYTDQIE 119

RESULT 12

ID 085150 PRELIMINARY; PRT; 148 AA.

AC 085150; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DR Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,

Wolf H., Modrow S.; "XXXSequence variability among different parvovirus B19 isolates.";

RT Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RL EMBL; Z70559; CAA94471.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 148 148

SQ SEQUENCE 148 AA; 16539 MW; 8B14ECF2459B308B CRC64;

Query Match 100.0%; Score 64; DB 12; Length 148;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 109 NKGTQOYTDQIE 120

DB 108 NKGTQOYTDQIE 119

RESULT 13

ID 085131 PRELIMINARY; PRT; 151 AA.

AC 085131; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DR Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,

Wolf H., Modrow S.; "Sequence variability among different parvovirus B19 isolates.";

RT J. Gen. Virol. 77:1781-1785(1996).

RL EMBL; Z70540; CAA94452.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 151 151

SQ SEQUENCE 151 AA; 16902 MW; 3CA74914B8E73A3E CRC64;

Query Match 100.0%; Score 64; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 109 NKGTQOYTDQIE 120

RESULT 14

ID 085177 PRELIMINARY; PRT; 151 AA.

AC 085177; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DR Viral protein 1 (Fragment).

GN VP1.

OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

NCBI_TaxID=10798;

[1] SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotzki A., Gigler A., Cassinotti P., Siegl G.,

Wolf H., Modrow S.; "XXXSequence variability among different parvovirus B19 isolates.";

RT Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RL EMBL; Z70586; CAA94499.1; -

DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1 1

FT NON_TER 151 151

SQ SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.0008;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
DB 109 NKGTQOYTDQIE 120

RESULT 15

085123 PRELIMINARY; PRT; 153 AA.
 AC 085123;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96332516; PubMed=8760426;
 RA Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "Sequence variability among different parvovirus B19 isolates."
 RL J. Gen. Virol. 77:1781-1785(1996).
 DR EMBL; 270532; CAA94444.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 153
 FT SEQUENCE 153 AA; 17131 MW; ECCA1F44020814EC CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 153;
 Best Local Similarity 100.0%; Pred. No. 0.00081;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

085123 PRELIMINARY; PRT; 157 AA.
 AC 085123;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270603; CAA94518.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 157
 FT SEQUENCE 157 AA; 17591 MW; 1D42191887FFCE03 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 157;
 Best Local Similarity 100.0%; Pred. No. 0.00083;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
 DB 110 NKGTOOYTDQIE 121

085135 PRELIMINARY; PRT; 162 AA.
 AC 085135;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)

01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemauer A., Von Poblotzki A., Gigler A., Caesinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXSequence variability among different parvovirus B19 isolates."
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; 270544; CAA94456.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1 162
 FT SEQUENCE 162 AA; 18005 MW; 3F51443566660F2 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 162;
 Best Local Similarity 100.0%; Pred. No. 0.00086;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
 DB 109 NKGTOOYTDQIE 120

0913X0 PRELIMINARY; PRT; 546 AA.
 AC 0913X0;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Iali;
 RA Hokynar K., Soderlund-Venermo M., Ranki A., Kiviniemi O., Partio E.K.,
 RA Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY044266; AAK95572.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 546 546
 FT SEQUENCE 546 AA; 59934 MW; 467BE468A67282E5 CRC64;
 SQ
 Query Match 100.0%; Score 64; DB 12; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
 DB 395 NKGTOOYTDQIE 406

0913W7 PRELIMINARY; PRT; 546 AA.
 AC 0913W7;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Major virus capsid protein VP2 (Fragment).
 GN VP2.
 OS Human parvovirus B19.

OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAM;
 RA Hokynar K., Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K., Hedman K.;
 RT "A new parvovirus B19 genotype persistent in skin."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AY044268; AAK55575.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON TER 546 546
 SQ SEQUENCE 546 AA; 59910 MW; BBIID23695E3FCD CRC64;

Query Match 100.0%; Score 64; DB 12; Length 546;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQIE 12
 |||||
 395 NKGTQOYTDQIE 406

RESULT 20
 O9PZS9 PRELIMINARY; PRT; 554 AA.
 ID O9PZS9;
 AC O9PZS9;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE VP2 capsid protein.
 GN VP.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HV;
 RA Gallinella G., Venturoli S.;
 RT "B19 Genome Sequence and Structure Analysis."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF162273; AAD46615.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60833 MW; 7901FEBD65A697B6 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
 |||||
 DB 395 NKGTQOYTDQIE 406

RESULT 21
 O90201 PRELIMINARY; PRT; 554 AA.
 ID O90201;
 AC O90201;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE VP2 structural protein (Capsid protein VP2) (Fragment).
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97081188; PubMed=8922470;
 RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the

RT VP1/VP2 gene from multiple isolates."
 RL J. Gen. Virol. 77:2767-2774(1996).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MI, and NB;
 RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
 RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NB;
 RX MEDLINE=90218047; PubMed=2157807;
 RA Umene K., Nunoe T.;

RT "The genome type of human parvovirus B19 strains isolated in Japan during 1981 differs from types detected in 1986 to 1987: a correlation between genome type and prevalence."
 RT J. Gen. Virol. 71:983-986(1990).

RL EMBL, U53595; AAB47453.1; -
 DR EMBL, U53596; AAB47455.1; -
 DR EMBL, U53593; AAB47449.1; -
 DR EMBL, U53594; AAB47451.1; -
 DR EMBL, U53597; AAB47457.1; -
 DR EMBL, U53600; AAB47463.1; -
 DR EMBL, U53598; AAB47465.1; -
 DR EMBL, U53599; AAB47461.1; -
 DR EMBL, AB030693; BAA90290.1; -
 DR EMBL, AB030673; BAA90268.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON TER 554 554
 SQ SEQUENCE 554 AA; 60853 MW; 98FE598F20CB66F CRC64;

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
 |||||
 DB 395 NKGTQOYTDQIE 406

RESULT 22
 O65790 PRELIMINARY; PRT; 554 AA.
 ID O65790;
 AC O65790;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VEN;
 RA Echevarria Mayo J.E., Erdman D.D.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL, U11358; AAA83559.1; -
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60798 MW; BA89F2B293BAE24 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQIE 12
 |||||
 DB 395 NKGTQOYTDQIE 406

RESULT 23
 Q912B7 PRELIMINARY; PRT; 554 AA.
 ID Q912B7
 AC Q912B7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE VP2 protein.
 GN VP2.
 OS Human erythrovirus V9.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=72197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V9;
 RA Nguyen Q.T.;
 RT "Molecular cloning and sequencing of a novel human erythrovirus genome: new species beside B19 in the genus Erythrovirus."
 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 EMBL; AJ249437; CAC80622.1; -.
 InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60885 MW; B4F338CDDA80F336 CRC64;
 SO

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 DB 395 NKGTQYTDQIE 406

RESULT 24
 Q9UGP7 PRELIMINARY; PRT; 554 AA.
 ID Q9UGP7
 AC Q9UGP7
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Capid protein VP2.
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rm;
 RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
 RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."
 Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB030694; BAA90293.1; -.
 InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60839 MW; 0DB958B33C735F64 CRC64;
 SO

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 DB 395 NKGTQYTDQIE 406

RESULT 25
 Q9WKL9 PRELIMINARY; PRT; 554 AA.
 ID Q9WKL9
 AC Q9WKL9
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Structural protein VP2.
 GN VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20014169; PubMed=10548133;
 RA Hemaier A., Beckenlehner K., Wolf H., Lang B., Modrow S.;
 RT "Acute parvovirus B19 infection in connection with a flare of systemic lupus erythematosus in a female patient."
 J. Clin. Virol. 14:73-77(1999).
 DR EMBL; AF13323; AAC99439.1; -.
 InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60913 MW; 98FB588FF205C66F CRC64;
 SO

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 DB 395 NKGTQYTDQIE 406

RESULT 26
 Q8JNS4 PRELIMINARY; PRT; 554 AA.
 ID Q8JNS4
 AC Q8JNS4
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE VP2.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D91.1;
 RX MEDLINE=22174902; PubMed=12186896;
 RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Merlier J.F., Garbarg-Chenon A.;
 RT "Genetic diversity within Human Erythroviruses: Identification of Three Genotypes."
 J. Virol. 76:9124-9134(2002).
 DR EMBL; AY083234; AAL91014.1; -.
 InterPro: IPR001403; Parvo_coat.
 DR Pfam: PF00740; Parvo_coat; 1.
 SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;
 SO

Query Match 100.0%; Score 64; DB 12; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQYTDQIE 12
 DB 395 NKGTQYTDQIE 406

RESULT 27
 Q9PZT8 PRELIMINARY; PRT; 760 AA.
 ID Q9PZT8
 AC Q9PZT8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE VP1/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat12;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161225; AAD45915.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 760
SQ SEQUENCE 760 AA; 83403 MW; EA6E0145E3A0E5A CRC64;

Query Match 100.0%; Score 64; DB 12; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
|||||
622 NKGTOOYTDQIE 633

```

```

RESULT 28
ID Q9PZU0 PRELIMINARY; PRT; 761 AA.
AC Q9PZU0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat11;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161223; AAD45910.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 761
SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 761;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
|||||
622 NKGTOOYTDQIE 633

RESULT 29
ID Q9PZT6 PRELIMINARY; PRT; 765 AA.
AC Q9PZT6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat13;

```

```

RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RA Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161225; AAD45915.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 765
SQ SEQUENCE 765 AA; 83998 MW; 89E2546086DCBB8 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 765;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
|||||
622 NKGTOOYTDQIE 633

```

```

RESULT 30
ID Q9PZT4 PRELIMINARY; PRT; 769 AA.
AC Q9PZT4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kat14;
RX MEDLINE=20191963; PubMed=10725428;
RA Hokynar K., Brunstein J., Soderlund-Venemo M., Kiviluoto O.,
RA Partio E.K., Kontinen Y., Hedman K.;
RT "Integrity and full coding sequence of B19 virus DNA persisting in
RT human synovial tissue.";
RL J. Gen. Virol. 81:1017-1025(2000).
DR EMBL; AF161226; AAD45917.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 769
SQ SEQUENCE 769 AA; 84578 MW; 0749D46B5CA7BB68 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 769;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOOYTDQIE 12
|||||
622 NKGTOOYTDQIE 633

RESULT 31
ID Q913X1 PRELIMINARY; PRT; 773 AA.
AC Q913X1;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Minor virus capsid protein VPI (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Labi;
RA Hokynar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,

```

RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95571.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85164 MW; D60CCAEF90B05378 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAM;
RA Hokyar K.; Soderlund-Venermo M., Ranki A., Kivluoto O., Partio E.K., Hedman K.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95571.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85140 MW; 2B86F345F9CEB50 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95571.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85140 MW; 2B86F345F9CEB50 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAM;
RA Hokyar K.; Soderlund-Venermo M., Ranki A., Kivluoto O., Partio E.K., Hedman K.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95571.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85140 MW; 2B86F345F9CEB50 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAM;
RA Hokyar K.; Soderlund-Venermo M., Ranki A., Kivluoto O., Partio E.K., Hedman K.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95571.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85140 MW; 2B86F345F9CEB50 CRC64;

RESULT 33

Q8JUN56 PRELIMINARY; PRT; 781 AA.
AC 08JUN56;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D91.1;
RX MEDLINE=22174902; PubMed=12186896;
RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Merlet J.F., Garbarg-Chenon A.;
RT "Genetic Diversity within Human Erythroviruses: Identification of Three Genotypes."
RT J. Virol. 76:9124-9134(2002).
DR EMBL; AY083234; AAL91013.1; -
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86142 MW; 4EB71FEDD41FC8F3 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DD CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQIE 12
ID 065789 PRELIMINARY; PRT; 781 AA.
AC 065789;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE VPI.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VENI;
RA Echavarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAB3558.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

RESULT 36
ID 085191 PRELIMINARY; PRT; 781 AA.
AC 085191;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Orf2 protein.
GN ORF2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Podolczki A., Giegler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.,
"XXXSequence variability among different parvovirus B19 isolates.";
Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
EMBL; Z70599; CA94513.1; -.
InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86026 MW; AA02577B63EDB2D CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
Db 622 NKGTOOYTDQIE 633
|||||
P90221 PRELIMINARY; PRT; 781 AA.
ID P90221;
AC P90221;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE VPI and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=KOR2;
MEDLINE=97081188; PubMed=8922470;
RT Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
"Genetic diversity of human parvovirus B19: sequence analysis of the
VPI/VP2 gene from multiple isolates.";
J. Gen. Virol. 77:2767-2774(1996).
RL EMBL; U38511; AAB4793.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; 9FA830083FEF1357 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
Db 622 NKGTOOYTDQIE 633
|||||
P89318 PRELIMINARY; PRT; 781 AA.
ID P89318;
AC P89318;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE VPI and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=JAP1;
RC MEDLINE=97081188; PubMed=8922470;
RX Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
"Genetic diversity of human parvovirus B19: sequence analysis of the
VPI/VP2 gene from multiple isolates.";
J. Gen. Virol. 77:2767-2774(1996).
RL EMBL; U38509; AAB4791.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86064 MW; 3AE65CE6906339 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
Db 622 NKGTOOYTDQIE 633
|||||
08JYE3 PRELIMINARY; PRT; 781 AA.
ID 08JYE3;
AC 08JYE3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 7.5 kDa protein.
OS Erythrovirus A6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=182494;
RN [1]
RP SEQUENCE FROM N.A.
RT Nguyen Q.T., Wong S., Brown K.E.;
"Identification and characterization of a second novel human
erythrovirus variant, A6.";
J. Virol. 76:100-106(2002).
RL EMBL; AY064475; AAL55418.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86277 MW; E73AB63E4551128 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQIE 12
Db 622 NKGTOOYTDQIE 633
|||||
P90224 PRELIMINARY; PRT; 781 AA.
ID P90224;
AC P90224;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE VPI and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxId=10798;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=BR21;
RC MEDLINE=97081188; PubMed=8922470;

RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
VP1/VP2 gene from multiple isolates";
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; U38546; AAB47801.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86039 MW; C3A29EB1DB8AD378 CRC64;

Query Match 100.0%; Score 64; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQIE 12
|||
Db 622 NKGTQOYTDQIE 633

Search completed: August 20, 2003, 09:23:58
Time : 34.886 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 34.0206 Seconds
(without alignments)
46.656 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKGTQYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Database: number of hits satisfying chosen parameters: 1107863

Maximum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	370	12	AA13406
2	55	100.0	543	12	AA13405
3	55	100.0	554	16	AAW08987
4	55	100.0	554	20	AAW23230
5	55	100.0	554	24	ABP57264
6	55	100.0	554	24	ABP57267
7	55	100.0	756	21	AAW71231
8	55	100.0	781	16	AAW08986
9	55	100.0	781	20	AAW23227

10	55	100.0	781	24	ABP57263
11	55	100.0	781	24	ABP57266
12	41	74.5	237	22	ABW60093
13	39	70.9	277	22	ABW64125
14	35	63.6	169	23	ABW54752
15	35	63.6	415	20	AAW06398
16	35	63.6	726	23	ABW77613
17	34	61.8	187	20	AAW14929
18	34	61.8	372	23	AAE25906
19	34	61.8	373	19	AAW40262
20	34	61.8	373	19	AAW40261
21	34	61.8	373	20	AAW42427
22	34	61.8	373	20	AAW42428
23	34	61.8	373	23	AAE25905
24	34	61.8	373	23	ABW76293
25	34	61.8	373	23	ABW76294
26	34	61.8	387	18	AAW08585
27	34	61.8	466	18	AAW08584
28	34	61.8	466	19	AAW40259
29	34	61.8	466	20	AAW42425
30	34	61.8	466	22	AAW00432
31	34	61.8	466	23	AAE25903
32	34	61.8	466	23	ABW76291
33	34	61.8	466	23	ABW92136
34	34	61.8	493	20	AAW23206
35	34	61.8	572	13	AAW26574
36	34	61.8	651	11	AAW06460
37	34	61.8	651	14	AAW33769
38	34	61.8	651	17	AAW06419
39	34	61.8	651	20	AAW23197
40	34	61.8	651	20	AAW23199
41	34	61.8	651	20	AAW23200
42	34	61.8	652	12	AAW14047
43	34	61.8	652	20	AAW23201
44	34	61.8	652	20	AAW23202
45	34	61.8	652	20	AAW23203

ALIGNMENTS

RESULT 1	AA13406	standard; Protein; 370 AA.
ID	AA13406	
AC	AA13406;	
XX		
DT	24-OCT-1991	(first entry)
XX		
DE	Parvo virus B19 PANSE.	
XX		
KW	Primer; PCR; PAPST; globulin.	
OS	Synthetic.	
XX		
PN	DE4003826-A.	
XX		
PD	14-AUG-1991.	
XX		
PF	08-FEB-1990;	90DE-4003826.
XX		
PR	08-FEB-1990;	90DE-4003826.
XX		
PA	(MIR-) MIKROGEN MOLEKULARB.	
XX		
PI	Soutschek E, Motz M;	
XX		
DR	WPI; 1991-246423/34.	
PT	Immunologically active parvo virus B19 peptide(s) - comprising	
PT	capsid protein VP1 or VP2 fragments, useful for antibody	
PT	detection or vaccination	
XX		

PS Claim 16; Page 10-11; 22pp; German.
 CC VP2 (AAR13405) and its fragments PANSE and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAR13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 CC
 XX
 SQ Sequence 370 AA;

Query Match: 100.0%; Score 55; DB 12; Length 370;
 Best Local Similarity 100.0%; Pred. No. 0.022; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

1 NKGTQOYTDO 10
 |||||
 231 NKGTQOYTDO 240

RESULT 2
 AAR13405
 ID AAR13405 standard; Protein; 543 AA.
 XX
 AC AAR13405;
 XX
 DT 24-OCT-1991 (first entry)
 XX
 DE Parvo virus B19 VP2.
 XX
 DE Primer; PCR; globulin; PANSE; PAPST.
 XX
 OS Synthetic.
 XX
 PN DE4003826-A.
 XX
 PD 14-AUG-1991.
 XX
 PF 08-FEB-1990; 90DE-4003826.
 XX
 PR 08-FEB-1990; 90DE-4003826.
 XX
 PA (MIKR-) MIKROGEN MOLEKULARB.
 XX
 YV Sautschek E, Motz M;
 DR WPI; 1991-246423/34.
 XX
 PT Immunologically active parvo virus B19 peptide(s) - comprising
 PT capsid protein VP1 or VP2 fragments, useful for antibody
 PT detection or vaccination
 XX
 PS Disclosure; Fig 2-6; 22pp; German.
 XX
 CC VP2 and its fragments PANSE (AAR13406) and PAPST (AAR13407) are useful
 CC as immunoassay reagents for detection of anti-B19 antibodies, e.g.
 CC for diagnosis of B19 infections, determining the immune status of
 CC pregnant woman, testing stored blood, or selecting positive donors
 CC for prodn. of B19 hyperimmune globulin preps.
 CC The VP2 fragments are expressed by plasmid pUC12VP1, which contains
 CC DNA sequences generated from viral DNA by PCR using the primers
 CC represented in AAR13159-63 in pairs.
 CC See also AAR13400-07 and AAR13414.
 CC
 XX
 SQ Sequence 543 AA;

Query Match: 100.0%; Score 55; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 0.034; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

OY 1 NKGTQOYTDO 10
 |||||
 Db 404 NKGTQOYTDO 413

RESULT 3
 AAM08987
 ID AAM08987 standard; Protein; 554 AA.
 XX
 AC AAM08987;
 XX
 DT 27-FEB-1997 (first entry)
 XX
 DE Human parvovirus VP-2 protein.
 XX
 DE Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KW non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KW erythblastemia; abortion; universal fetal hydrops; liver disease;
 KW haemorrhagic fever; rheumatism; detection; IgG antibody.
 XX
 OS Human parvovirus.
 XX
 PN JP07147986-A.
 XX
 PD 13-JUN-1995.
 XX
 PF 24-SEP-1992; 92JP-0281017.
 XX
 PR 24-SEP-1992; 92JP-0281017.
 XX
 PA (DENK-) DENKA SEIKEN KK.
 PA (ELED-) DENKI KAGAKU KOGYO KK.
 XX
 DR WPI; 1995-242756/32.
 DR N-PSDB; AAT49535.
 XX
 PT Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.
 XX
 PS Claim 3; Page 7-9; 38pp; English.
 XX
 CC The sequences given in AAM08986 represent the parvovirus structural
 CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.
 XX
 SQ Sequence 554 AA;

Query Match: 100.0%; Score 55; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.035; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

OY 1 NKGTQOYTDO 10
 |||||
 Db 395 NKGTQOYTDO 404

RESULT 4
 AAY23230
 ID AAY23230 standard; Protein; 554 AA.
 XX
 AC AAY23230;
 XX
 DT 26-AUG-1999 (first entry)
 XX
 DE Erythrovirus V9 VP2 protein.
 XX
 KW Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KW erythrovirus screening; typing; immunoassay; VP2 protein.

```
XX OS Erythrovirus.
XX XX FR2771751-A1.
XX PD 04-JUN-1999.
XX PF 03-DEC-1997; 97FR-0015197.
XX PR 03-DEC-1997; 97FR-0015197.
XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
XX P1 Auguste V, Garbary CA, Nguyen QT;
XX DR WPI; 1999-349543/30.
XX DR N-PSDB; ABX81586.
XX PT Erythrovirus V9 and its nucleic acid sequences - can be used in the
XX PT diagnosis of its infections
XX PT Claim 19; Page 57-58; 80pp; French.
XX CC The present sequence represents an erythrovirus V9 protein.
XX CC Probes and primers derived from erythrovirus V9 polynucleotide
XX CC sequences (ABX81580) can be used for differential diagnosis of
XX CC erythrovirus (parvovirus) infections by a combination of
XX CC amplification and hybridisation assay. The probes can also be
XX CC used to assess susceptibility to erythrovirus infection and
XX CC for erythrovirus screening and typing. The antibodies can be
XX CC used in immunoassays for diagnosis of erythrovirus V9 infections.
XX SQ Sequence 554 AA;
OY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404
Query Match 100.0%; Score 55; DB 20; Length 554;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404
RESULT 5
ABP57264
ID ABP57264 standard; Protein; 554 AA.
XX AC ABP57264;
XX AC 22-APR-2003 (first entry)
XX DE Human parvovirus B19 clone B1-VP2 amino acid sequence SEQ ID NO:29.
XX KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX OS Human parvovirus B19.
XX PN WO2003002753-A2.
XX PD 09-JAN-2003.
XX PF 28-JUN-2002; 2002WO-US20684.
XX PR 28-JUN-2001; 2001US-302077P.
XX PR 19-MAR-2002; 2002US-365956P.
XX PR 29-MAR-2002; 2002US-369224P.
XX PA (CHIR ) CHIRON CORP.
XX PI Pichuanes S, Shyamala V;
XX DR WPI; 2003-201510/19.
XX DR N-PSDB; ABZ59574.
```

```
XX XX Detecting a human parvovirus B19 infection in a biological sample to
XX PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
XX PT acid with a primer complementary to the 3'-terminal portion of the RNA
XX PT target sequence -
XX XX Example 4; Fig 7B; 148pp; English.
XX XX The present invention describes a method for detecting a human parvovirus
XX CC B19 infection in a biological sample. The method comprises reacting the
XX CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
XX CC consisting of a first primer containing a complexing sequence
XX CC sufficiently complementary to the 3'-terminal portion of the RNA target
XX CC sequence to complex with. Also described: (1) amplifying a target
XX CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
XX CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
XX CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
XX CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
XX CC consisting of a promoter region recognised by a DNA-dependent RNA
XX CC polymerase operably linked to a human parvovirus B19-specific complexing
XX CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
XX CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
XX CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
XX CC oligonucleotide primer of (4), and instructions for conducting the
XX CC diagnostic test. The method is useful for detecting parvovirus infection
XX CC in a biological sample, such as in blood products, to prevent
XX CC transmission of the virus through blood and plasma derivatives or by
XX CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
XX CC represent sequences used in the exemplification of the present invention.
XX SQ Sequence 554 AA;
OY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404
Query Match 100.0%; Score 55; DB 24; Length 554;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404
RESULT 6
ABP57267
ID ABP57267 standard; Protein; 554 AA.
XX AC ABP57267;
XX AC 22-APR-2003 (first entry)
XX DE Human parvovirus B19 clone B6-VP2 amino acid sequence SEQ ID NO:35.
XX KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
XX OS Human parvovirus B19.
XX PN WO2003002753-A2.
XX PD 09-JAN-2003.
XX PF 28-JUN-2002; 2002WO-US20684.
XX PR 28-JUN-2001; 2001US-302077P.
XX PR 19-MAR-2002; 2002US-365956P.
XX PR 29-MAR-2002; 2002US-369224P.
XX PA (CHIR ) CHIRON CORP.
XX PI Pichuanes S, Shyamala V;
XX DR WPI; 2003-201510/19.
XX DR N-PSDB; ABZ59577.
XX PT Detecting a human parvovirus B19 infection in a biological sample to
```

PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -

PS Example 4; Fig 10B; 148pp; English.

CC The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see AB259549 to AB259569, and AB259604 to
 CC AB259629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see AB259570 and AB259571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. AB259549 to AB259634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.

SO Sequence 554 AA:

Query Match 100.0%; Score 55; DB 24; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
 |||||
 DB 395 NKGTOOYTDQ 404

RESULT 7

ID AAV71231 standard; Protein; 756 AA.

AC AAV71231;

DT 08-SEP-2000 (first entry)

XX Capsid protein encoded by AAV2/human parvovirus B19 chimeric vector.

KM Adeno-associated virus; AAV2; human parvovirus B19; chimeric;
 KM recombinant parvoviral vector; cellular tropism; cap protein;
 KM capsid; gene delivery; gene therapy; VP1; VP2; VP3.

OS Chimeric - Adeno associated virus serotype 2.

OS Chimeric - Human parvovirus B19.

PN WO200028004-A1.

PD 18-MAY-2000.

PF 10-NOV-1999; 99WO-US26505.

PR 10-NOV-1998; 98US-0107840.

PR 10-MAR-1999; 99US-0123651.

PA (UNNC-) UNIV NORTH CAROLINA.

PI Rabinowitz JE, Samulecki RJ, Xiao W;

DR MPI: 2000-376523/32.

DR N-PSDB; AAD00833.

PT Recombinant parvoviral vectors with altered packaging, tropisms and

PT immunogenic properties, useful in gene therapy protocols -

PS Example 21; Page 142; 153pp; English.

CC The patent discloses modified parvovirus vectors with advantageous
 CC antigenic properties, packaging capabilities and cellular tropisms.
 CC These vectors can be used in standard recombinant DNA protocols e.g. gene
 CC therapy for delivering nucleic acids to cells.
 CC The present sequence is a capsid protein encoded by an
 CC adeno-associated virus serotype 2 (AAV2)/human parvovirus B19 chimeric
 CC vector. This vector encodes AAV2 VP1 and VP2 capsid proteins
 CC and human parvovirus B19 VP2 protein. The chimeric vector was
 CC constructed by replacing the VP3 major cap protein of AAV2 with B19's
 CC VP2. Recombinant parvovirus comprising the chimeric capsid is useful
 CC for gene delivery.

SO Sequence 756 AA:

Query Match 100.0%; Score 55; DB 21; Length 756;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
 |||||
 DB 597 NKGTOOYTDQ 606

RESULT 8

ID AAM08986 standard; Protein; 781 AA.

AC AAM08986;

DT 27-FEB-1997 (first entry)

XX Human parvovirus VP-1 protein.

KM Human; parvovirus genome; structural gene; VP-1; VP2; arthritis;
 KM non-structural protein; NS; diagnosis; vaccine; parvoviral disease;
 KM erythraemia; abortion; universal fetal hydrops; liver disease;
 KM haemorrhagic fever; rheumatism; detection; IgG antibody.

OS Human parvovirus.

XX Key Location/Qualifiers

FT Misc-difference 19 /label= Gly, Val

FT Misc-difference 61 /label= Asn, Asp

FT Misc-difference 220 /label= His, Asn

PN JP07147986-A.

PD 13-JUN-1995.

PF 24-SEP-1992; 92JP-0281017.

PR 24-SEP-1992; 92JP-0281017.

PA (DENK-) DENKA SEIKEN KK.

PA (ELED) DENKI KAGAKU KOGYO KK.

DR MPI: 1995-242756/32.

DR N-PSDB; AAT49535.

PT Human parvovirus gene coding for a polypeptide - useful for
 PT developing vaccines against parvoviral diseases such as
 PT erythroblastemia, haemorrhagic fever, etc.

PS Claim 2; Page 5-7; 38pp; English.

CC The sequences given in AAM08986 represent the parvovirus structural

CC proteins, VP-1 and VP2, and the non-structural protein, NS. The
 CC genomic fragment encoding these proteins may be used for the
 CC diagnosis and development of vaccines for parvoviral diseases including
 CC erythblastemia, abortion, universal fetal hydrops, liver diseases,
 CC haemorrhagic fever, arthritis and rheumatism. The VP-1 and VP-2
 CC proteins may be used to detect parvovirus IgG antibodies.

XX Sequence 781 AA;

Query Match 100.0%; Score 55; DB 16; Length 781;

Best Local Similarity 100.0%; Pred. No. 0.051; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
 |||||
 DB 622 NKGTOQYTDQ 631

RESULT 9

AAV23227 standard; Protein; 781 AA.

AAV23227;

26-AUG-1999 (first entry)

Erythrovirus V9 VP1 protein.

Erythrovirus V9; differential diagnosis; parvovirus; infection;

Erythrovirus screening; typing; immunoassay; VP1 protein.

Erythrovirus.

FR2771751-A1.

04-JUN-1999.

03-DEC-1997; 97FR-0015197.

03-DEC-1997; 97FR-0015197.

(ASSISTANCE PUBLIQUE HOPITAUX PARIS.

Auguste V, Garbarg CA, Nguyen QT;

WPI; 1999-349543/30.

N-PSDB; AAX81583.

Erythrovirus V9 and its nucleic acid sequences - can be used in the
 diagnosis of its infections

Claim 19; Page 50-52; 80pp; French.

XX The present sequence represents an erythrovirus V9 protein.
 CC probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.

XX Sequence 781 AA;

Query Match 100.0%; Score 55; DB 20; Length 781;

Best Local Similarity 100.0%; Pred. No. 0.051; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
 |||||
 DB 622 NKGTOQYTDQ 631

RESULT 10

ABP57263 standard; Protein; 781 AA.

ABP57263;

22-APR-2003 (first entry)

Human parvovirus B19 clone B1-VP1 amino acid sequence SEQ ID NO:27.

Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.

Human parvovirus B19.

MO2003002753-A2.

09-JAN-2003.

28-JUN-2002; 2002WO-US20684.

28-JUN-2001; 2001US-302077P.

19-MAR-2002; 2002US-365956P.

29-MAR-2002; 2002US-369224P.

(CHIR) CHIRON CORP.

Pichuanes S, Shyamala V;

WPI; 2003-201510/19.

N-PSDB; ABZ59573.

Detecting a human parvovirus B19 infection in a biological sample to
 prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 acid with a primer complementary to the 3'-terminal portion of the RNA
 target sequence

Example 4; Fig 6B; 148pp; English.

XX The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC sequence to complex with. Also described: (1) amplifying a target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 55; DB 24; Length 781;

Best Local Similarity 100.0%; Pred. No. 0.051; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
 |||||
 DB 622 NKGTOQYTDQ 631

RESULT 11

ABP57266

ID ABP57266 standard; Protein; 781 AA.
 XX
 AC ABP57266;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Human parvovirus B19 clone B6-VPI amino acid sequence SEQ ID NO:33.
 XX
 KM Human parvovirus B19; parvovirus B19; infection; virus; blood; plasma.
 XX
 OS Human parvovirus B19.
 XX
 PN WO2003002753-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WO-US20684.
 XX
 PP 28-JUN-2001; 2001US-302077P.
 XX
 PR 19-MAR-2002; 2002US-365956P.
 XX
 PS 29-MAR-2002; 2002US-369224P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Pichuanes S, Shyamala V;
 XX
 DR MPI: 2003-201510/19.
 XX
 DR N-PSDB; ABZ59576.
 XX

PT Detecting a human parvovirus B19 infection in a biological sample to
 PT prevent viral transmission, comprises reacting a parvovirus B19 nucleic
 PT acid with a primer complementary to the 3'-terminal portion of the RNA
 PT target sequence -

XX Example 4; Fig 9B; 148bp; English.

XX The present invention describes a method for detecting a human parvovirus
 CC B19 infection in a biological sample. The method comprises reacting the
 CC isolated parvovirus B19 nucleic acid with a first oligonucleotide
 CC consisting of a first primer containing a complexing sequence
 CC sufficiently complementary to the 3'-terminal portion of the RNA target
 CC parvovirus B19 nucleotide sequence; (2) a polynucleotide comprising one
 CC of 47 700 base pair sequences (see ABZ59549 to ABZ59569, and ABZ59604 to
 CC ABZ59629); (3) a polynucleotide comprising either of 2 4678 base pair
 CC sequences (see ABZ59570 and ABZ59571); (4) an oligonucleotide primer
 CC consisting of a promoter region recognised by a DNA-dependent RNA
 CC polymerase operably linked to a human parvovirus B19-specific complexing
 CC sequence of 10-75 nucleotides; (5) an oligonucleotide probe comprising a
 CC parvovirus B19-specific hybridising sequence of 10-50 nucleotides linked
 CC to an acridinium ester label; and (6) a diagnostic test kit comprising an
 CC oligonucleotide primer of (4), and instructions for conducting the
 CC diagnostic test. The method is useful for detecting parvovirus infection
 CC in a biological sample, such as in blood products, to prevent
 CC transmission of the virus through blood and plasma derivatives or by
 CC close personal contact. ABZ59549 to ABZ59634 and ABP57262 to ABP57267
 CC represent sequences used in the exemplification of the present invention.
 XX

XX Sequence 781 AA;

Query Match 100.0%; Score 55; DB 24; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDO 10
 |||||
 DB 622 NKGTOOYTDO 631

RESULT 12
 ABB60093
 ID ABB60093 standard; Protein; 237 AA.
 XX

AC ABB60093;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7071.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PP 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PS 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR MPI: 2001-656860/75.
 XX
 DR N-PSDB; ABL04196.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Disclosure; SEQ ID NO 7071; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL016176-ABL030511), expressed DNA
 CC sequences (ABL016176-ABL016175) and the encoded proteins
 CC (ABZ57737-ABZ72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 237 AA;

Query Match 74.5%; Score 41; DB 22; Length 237;
 Best Local Similarity 70.0%; Pred. No. 7;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOYTDO 10
 |||||
 DB 30 NKGTOOYTDO 39

RESULT 13
 ABB64125
 ID ABB64125 standard; Protein; 277 AA.
 XX

AC ABB64125;
 XX
 DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster polypeptide SEQ ID NO 19167.

KM Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.
 XX

PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL08228.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 19167; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
PS capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (AB57737-AB872072).
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 277 AA;
SQ

Query Match 70.9%; Score 39; DB 22; Length 277;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
|:|:|:|:|:
Db 247 NRGTRQYEDD 256

RESULT 14
AB54752
ID AB54752 standard; Protein; 169 AA.
XX
AC AB54752;
XX
XX 16-MAY-2002 (first entry)
XX
DE Lactococcus lactis protein p1339.
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
XX Lactococcus lactis IL1403.
XX
XX FR2807446-A1.
XX
XX 12-OCT-2001.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX 11-APR-2000; 2000FR-0004630.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
XX WPI; 2002-043418/06.
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -

XX
XX Claim 6; SEQ ID No 1454; 2504pp; French.
XX
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (AB53300-AB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 169 AA;
SQ

Query Match 63.6%; Score 35; DB 23; Length 169;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
|:|:|:|:|:
Db 60 NNSSQYADQ 69

RESULT 15
AA06398
ID AA06398 standard; Protein; 415 AA.
XX
AC AA06398;
XX
XX 20-SEP-1999 (first entry)
XX
XX Bacillus subtilis metalloprotease YmfH.
XX
XX Metalloprotease; protease; YmfH; detergent; surfactant; cleaning;
XX textile; feedstuff; animal feed; host cell.
XX
XX Bacillus subtilis.
XX
XX Key Location/Qualifiers
FH FT 55..59
FT Domain /note= "catalytic domain HXXEH motif"
XX
XX WO9333960-A2.
XX
XX 08-JUL-1999.
XX
XX 17-DEC-1998; 98WO-US27040.
XX
XX 30-DEC-1997; 97GB-0027471.
XX
XX (GENV) GENENCOR INT INC.
XX
XX Estell DA;
XX
XX WPI; 1999-419100/35.
XX
XX N-PSDB; AAX59344.
XX
XX A Bacillus subtilis metalloprotease, designated YmfH, useful in
PT cleaning compositions, animal feed and for treating textiles
XX
XX Claim 9; Fig 1A-F; 32pp; English.
XX
XX The present sequence represents a novel metalloprotease (MP),
CC designated YmfH, of Bacillus subtilis. YmfH DNA (see AAX59344) was
CC identified via a BLAST search of B. subtilis genomic DNA. The
CC deduced protein sequence shows identity to the MP p1339 of
CC Bacillus coli, including the presence of the catalytic domain
CC HXXEH+5. An expression vector including YmfH DNA and a host cell
CC comprising the vector are claimed. Also claimed are a cleaning
CC composition, an animal feed and a composition for the treatment of

CC a textile, all comprising YmfH. Gram positive microorganisms
 CC having a mutation or deletion of all or part of YmfH DNA are used
 CC as host cells for expression of a homologous or heterologous
 CC protein, such as a hormone, growth factor, cytokine or enzyme,
 CC especially a protease, carbohydrase, lipase, isomerase, oxidase,
 CC reductase, transferase, kinase or phosphatase (all claimed). Also
 CC claimed is a method for detecting a Gram positive microorganism MP
 CC using a probe comprising all or part of the YmfH DNA.

SO Sequence 415 AA;

Query Match 63.6%; Score 35; DB 20; Length 415;
 Best Local Similarity 60.0%; Pred. No. 2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOYTDO 10
 DB 208 NKGKPYTDO 217

ABBT 16
 ABB77613

ID ABB77613 standard; Protein; 726 AA.

AC ABB77613;

DT 01-JUL-2002 (first entry)

DE AMEPV NTPase (AMV087).

KW AMEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
 genetic deficiency disorder; NTPase.

OS Amsacta moorei entomopoxvirus.

PN WO200212526-A2.

PD 14-FEB-2002.

PF 10-AUG-2001; 2001WO-US25287.

PR 10-AUG-2000; 2000US-224479P.

PR 14-SEP-2000; 2000US-0662254.

PA (UYFL) UNIV FLORIDA.

PT Moyer RW, Li Y, Bawden AL;

WI; 2002-227161/28.

N-PSDB; ABL56216.

PT Novel recombinant entomopox virus vector useful for delivering
 PT polynucleotide encoding protein to vertebrate cell, comprises
 PT polynucleotide encoding protein operably linked with heterologous
 PT promoter sequence

Claim 75; Page 263-266; 326pp; English.

CC The invention relates to a recombinant entomopox virus (EPV) vector,
 CC comprising a polynucleotide encoding a protein operably linked with a
 CC heterologous promoter sequence. The invention also concerns methods for
 CC providing gene therapy for genetic deficiency disorders. Vectors of the
 CC invention are useful for delivering a polynucleotide encoding a protein
 CC to a vertebrate cell preferably a mammalian cell, such as a human cell.
 CC The vector is introduced into the vertebrate cell by infection in a viral
 CC particle, or by transfection, transduction, or injection either in vitro
 CC or in vivo. The vector is useful for the delivery and expression of
 CC biologically useful proteins in gene therapy protocols, and for
 CC delivering large DNA segments for engineering of vertebrate cells.
 CC Polynucleotides of the invention have applications in techniques such as
 CC their use as insertion sites for foreign genes of interest, hybridisation
 CC probes, for chromosome and gene mapping, in PCR technologies, and in the
 CC production of sense or antisense nucleic acids. Vectors of the invention

CC provide for stable integration and expression of heterologous DNA in host
 CC cells, and are adapted for accepting large heterologous polynucleotide
 CC inserts which can be delivered in an infected or transformed cell and
 CC expressed in a stable fraction. The current sequence represents an
 CC amsacta moorei entomopoxvirus (AMEPV) NTPase (AMV087).

SO Sequence 726 AA;

Query Match 63.6%; Score 35; DB 23; Length 726;
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOOYTDO 10
 DB 598 NKNKQYTRQ 607

RESULT 17
 AAY14929

ID AAY14929 standard; protein; 187 AA.

AC AAY14929;

DT 25-OCT-1999 (first entry)

DE Amino acid sequence of M. vaccae antigen GV-41B.

KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 dendritic cell maturation; infectious disease; immune disorder; cancer;

KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;

KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 squamous cell carcinoma; melanoma.

OS Mycobacterium vaccae.

PN WO9932634-A2.

PD 01-JUL-1999.

PF 23-DEC-1998; 98WO-NZ00189.

PR 04-DEC-1998; 98US-0205426.

PR 23-DEC-1997; 97US-0996624.

PR 23-DEC-1997; 97US-0997080.

PR 23-DEC-1997; 97US-0997362.

PR 11-JUN-1998; 98US-0095855.

PR 17-SEP-1998; 98US-0156181.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PT Prestidge RL, Skinner MA, Tan P, Visser ES, Watson U;

WI; 1999-430163/36.

N-PSDB; AAZ11394.

PT Enhancing immune response to an antigen

Claim 1; Page 240; 243pp; English.

CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as

CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma.

XX

Sequence 187 AA;

Query Match 61.8%; Score 34; DB 20; Length 187;

Best Local Similarity 60.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOQYTD 10

DB 161 DKSTHQYTNQ 170

RESULT 18

AAE25906

AAE25906;

15-NOV-2002 (first entry)

Human chitinase protein fragment.

Human; chitinase; chitin-binding fragment; therapy; fungal infection;
fungicide; enzyme.

Homo sapiens.

US6399571-B1.

04-JUN-2002.

12-MAR-1999; 99US-0267574.

12-MAR-1998; 98US-0039198.

(ICOS-) ICOS CORP.

Gray PW, Tjoelker LW;

WPI; 2002-606631/65.

New chitinase derived polypeptides used for treating fungal infections
and for detecting and binding chitin -

Example 3; Column 45-48; 32pp; English.

The invention relates to chitinase derived polypeptide, chitin-binding
fragments of human chitinase, fragment analogues, purified and isolated
polynucleotides encoding such fragments. The invention is used for
treating fungal infections and for detecting and binding chitin. The
present sequence is human chitinase protein fragment.

Sequence 372 AA;

Query Match 61.8%; Score 34; DB 23; Length 372;

Best Local Similarity 66.7%; Pred. No. 2.7e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOQYTD 9

DB 79 NFGTQKFTD 87

RESULT 19

AAW40262

AAW40262;

15-JUN-1998 (first entry)

Human chitinase protein analogue.

Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
Gaucher's disease; transgenic; detection; hybridisation; antifungal;
rheumatoid arthritis; overexpression; extracellular matrix.

Homo sapiens.
Synthetic.

Key Location/Qualifiers
FT Misc-difference 370
FT /label= P3705
FT /note= "Wild type Pro is replaced by Ser"

W09747752-A1.

18-DEC-1997.

16-JUN-1997; 97WO-US10460.

14-JUN-1996; 96US-0663618.

(ICOS-) ICOS CORP.

Gray PW;

WPI; 1998-052316/05.

Nucleic acids encoding human chitinase - useful as antifungal
agents, especially in combination with other antifungals

Claim 29; Page 49-50; 63pp; English.

This sequence encodes a fragment of a novel human chitinase protein
analogue in which a proline at position 370 of the wild type protein
(see AAW40261) is substituted with a serine and the C-terminal 72
residues of the mature protein are deleted. These residues are not
critical to chitinase enzymatic activity. Chitinases are useful for
treating or preventing fungal infection and as immunogens for generating
antibodies which are used to purify, detect and quantify chitinases, e.g.
for diagnosis of Gaucher's disease. The nucleic acid sequence of the
chitinase is also useful as a probe to identify and isolate genomic DNA
encoding chitinases or similar proteins, or cells expressing them or to
generate transgenic ('knockout') rodents. It can also be used in
hybridisation assays and to detect genetic alterations in the chitinase
gene related to disease. Agents that inhibit this protein may be useful
in treatment of Gaucher's disease and rheumatoid arthritis, where
overexpression of the protein can damage the extracellular matrix.
Chitinase also improves the activity of other antifungal agents and may
allow a reduction in the dose of such agents, and thus of their side
effects.

Sequence 373 AA;

Query Match 61.8%; Score 34; DB 19; Length 373;

Best Local Similarity 66.7%; Pred. No. 2.7e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOQYTD 9

DB 79 NFGTQKFTD 87

RESULT 20

AAW40261

AAW40261;

15-JUN-1998 (first entry)

Human chitinase protein fragment.

XX		Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KV		Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW		rheumatoid arthritis; overexpression; extracellular matrix.
OS	Homo sapiens.	
XX		
PN	WO9747752-A1.	
XX		
PD	18-DEC-1997.	
XX		
PF	16-JUN-1997; 97MO-US10460.	
XX		
PR	14-JUN-1996; 96US-0663618.	
XX		
PA	(ICOS-) ICOS CORP.	
PI		
PJ	Gray PW;	
XX		
WPI:	1998-052316/05.	
N-PSDB:	AAVL0435.	
XX		
PT	Nucleic acids encoding human chitinase - useful as antifungal	
PI	agents, especially in combination with other antifungals	
XX		
PS	Claim 23; Page 48-49; 63pp; English.	
XX		
CC	This sequence encodes a fragment of a novel human chitinase protein	
CC	which lacks the C-terminal 72 residues of the mature protein. These	
CC	residues are not critical to chitinase enzymatic activity. Chitinaes are	
CC	useful for treating or preventing fungal infection and as immunogens for	
CC	generating antibodies which are used to purify, detect and quantify	
CC	chitinaes, e.g. for diagnosis of Gaucher's disease. The nucleic acid	
CC	sequence of the chitinase is also useful as a probe to identify and	
CC	isolate genomic DNA encoding chitinaes or similar proteins, or cells	
CC	expressing them or to generate transgenic ('knockout') rodents. It can	
CC	also be used in hybridisation assays and to detect genetic alterations	
CC	in the chitinase gene related to disease. Agents that inhibit this	
CC	protein may be useful in treatment of Gaucher's disease and rheumatoid	
CC	arthritis, where overexpression of the protein can damage the	
CC	extracellular matrix. Chitinase also improves the activity of other	
CC	antifungal agents and may allow a reduction in the dose of such agents,	
CC	and thus of their side effects.	
XX		
SQ	Sequence 373 AA;	
Query Match	61.8%; Score 34; DB 19; Length 373;	
Local Similarity	66.7%; Pred. No. 2.7e+02;	
Conservative	2; Mismatches 1; Indels 0; Gaps 0;	
NCBI Query Match	61.8%; Score 34; DB 19; Length 373;	
Local Similarity	66.7%; Pred. No. 2.7e+02;	
Conservative	2; Mismatches 1; Indels 0; Gaps 0;	
Y	1 NKGTGQYTD 9	
Db	79 NFGTOKFTD 87	
RESULT 21		
AAAY42427		
ID	AAAY42427 standard; Protein; 373 AA.	
XX		
AC	AAAY42427;	
XX		
DT	10-DEC-1999 (first entry)	
XX		
DE	Clone of the C-terminal fragment of human chitinase.	
XX		
XX	chitin; fungal infection; immunocompromised; AIDS; chemotherapy;	
KV	organ transplant; parasite; chitin-binding; allele; vector;	
KW	truncated protein; bacterial expression.	
XX		
OS	Synthetic.	
XX		
SS	Homo sapiens.	
XX		
XX	WO9946390-A1	

XX 16-SEP-1999.
PD
XX
PD 12-MAR-1999; 99WO-US05343.
XX
PF 12-MAR-1998; 98US-0039198.
XX
PR 12-MAR-1998; 98US-0039198.
XX
XX
PA (ICOS-) ICOS CORP.
XX
XX Gray PM, Tjoelker LW;
PI WPI, 1999-551417/46.
XX
XX Novel chitin-binding fragments of human chitinase used to treat fungal
PT infections in animals
XX
XX
PS Example 5; Page 66-67; 83pp; English.
XX
XX This is the amino acid sequence of the C-terminal fragment of human
CC chitinase, which can be expressed in bacterial cells. The fragment was
CC prepared by amplifying the MO-218 plasmid with primers 218-1 (AAZ21855)
CC and T-END (AAZ21856).
CC
CC Chitinase fragments can be used to screen for proteins or other
CC molecules that specifically bind to the chitin-binding domain of human
CC chitinase or that modulate its activity. These compounds are useful for
CC immunization, as well as for purifying chitinase, as well as for
CC detection and quantification of chitinase. Polynucleotide fragments of
CC the invention are useful as a source of probes and primers, and to
CC express the proteins recombinantly. The chitinase fragments, when
CC conjugated to antifungal compounds, are used to treat animals,
CC especially humans, infected with chitin-containing parasites such as
CC fungi. Fungal infection treated include candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
CC sporotrichosis, and dermatophytoses.
CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
CC protein for treating infections, especially fungal infections, is
CC problematic. In view of the increasing incidents of life-threatening
CC fungal infection in e.g. immunocompromised individuals, there exists a
CC need for identifying new compounds for treating fungal infection. The
CC chitin-binding fragments of the present invention provide this need.
XX
XX
SQ Sequence 373 AA;

Query Match 61.8%; Score 34; DB 20; Length 373;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 1 NKGTOOYTD 9
| | | | | | | | | |
Db 79 NFGTOKEPTD 87

RESULT 22
AA42428
ID AAY42428 standard; Protein; 373 AA.
XX
XX AAY42428;
AC
XX
XX 10-DEC-1999 (first entry)
DT
XX
XX Chitinase amino acid fragment from plasmid PARACB.
DE
XX
XX Chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; bacterial expression.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX W09946390-A1.
XX
XX 16-SEP-1999.
PD

XX 12-MAR-1999; 99WO-US05343.
 PF 12-MAR-1998; 98US-0039198.
 PR (ICOS-) ICOS CORP.
 XX Gray PW, Tjoelker LW;
 PI WPI, 1999-551417/46.
 DR Novel chitin-binding fragments of human chitinase used to treat fungal
 PT infections in animals -
 PT Example 3; Page 67-68; 83pp; English.
 PS This is the chitinase fragment generated from the modified MO-218
 CC sequence incorporated into the pParC8 plasmid. This chitinase fragment
 CC has a serine at mature amino acid position 370 instead of the proline
 CC encoded by MO-218.
 CC Chitinase fragments can be used to screen for proteins or other
 CC molecules that specifically bind to the chitin-binding domain of human
 CC chitinase or that modulate its activity. These compounds are useful for
 CC immunization, as well as for purifying chitinase, as well as for
 CC detection and quantification of chitinase. Polynucleotide fragments of
 CC the invention are useful as a source of probes and primers, and to
 CC express the proteins recombinantly. The chitinase fragments, when
 CC conjugated to antifungal compounds, are used to treat animals,
 CC especially humans, infected with chitin-containing parasites such as
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 CC sporotrichosis, and dermatophytoses.
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 CC protein for treating infections, especially fungal infections, is
 CC problematic. In view of the increasing incidence of life-threatening
 CC fungal infection in e.g. immunocompromised individuals, there exists a
 CC need for identifying new compounds for treating fungal infection. The
 CC chitin-binding fragments of the present invention provide this need.
 XX Sequence 373 AA;
 SO
 Query Match 61.8%; Score 34; DB 20; Length 373;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTQGYTD 9
 79 NFGTQKFTD 87
 RESULT 23
 AAE25905
 ID AAE25905 standard; Protein; 373 AA.
 AC AAE25905;
 XX 15-NOV-2002 (first entry)
 DT Human chitinase-related protein.
 XX Human chitinase-related protein.
 DE Human chitinase; chitin-binding fragment; therapy; fungal infection;
 KW fungicide; enzyme.
 XX Homo sapiens.
 OS Homo sapiens.
 XX US6399571-B1.
 PN 04-JUN-2002.
 PD 12-MAR-1999; 99US-0267574.
 PF 12-MAR-1998; 98US-0039198.
 PR 12-MAR-1998; 98US-0039198.

XX (ICOS-) ICOS CORP.
 PA Gray PW, Tjoelker LW;
 PI WPI, 2002-606631/65.
 XX WPI, 2002-606631/65.
 DR New chitinase derived polypeptides used for treating fungal infections
 PT and for detecting and binding chitin -
 PT Example 5; Column 43-46; 32pp; English.
 PS The invention relates to chitinase derived polypeptide, chitin-binding
 CC fragments of human chitinase, fragment analogues, purified and isolated
 CC polynucleotides encoding such fragments. The invention is used for
 CC treating fungal infections and for detecting and binding chitin. The
 CC present sequence is human chitinase-related protein.
 XX Sequence 373 AA;
 SO
 Query Match 61.8%; Score 34; DB 23; Length 373;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTQGYTD 9
 79 NFGTQKFTD 87
 Db
 RESULT 24
 ABB76293
 ID ABB76293 standard; Protein; 373 AA.
 AC ABB76293;
 XX 12-AUG-2002 (first entry)
 DT Human chitinase truncated polypeptide.
 XX
 DE Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis.
 XX Homo sapiens.
 OS Homo sapiens.
 XX US6372212-B1.
 PN 16-APR-2002.
 PD 16-JUN-1997; 97US-0877559.
 PF 14-JUN-1996; 96US-0663618.
 PR (ICOS-) ICOS CORP.
 XX Gray PW;
 PI WPI, 2002-442449/47.
 DR Co-administering chitinase to improve the effectiveness of fungicidal
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 XX Example 5; Column 49-42; 26pp; English.
 PS The present sequence is the protein sequence of a human mature
 CC chitinase truncated polypeptide, in which the 72 C-terminal amino
 CC acids of the native mature protein (see ABB76291) are deleted.
 CC The coding region for the polypeptide was obtained by PCR from
 CC chitinase MO-218 cDNA (see AB157380), and the polypeptide was
 CC produced in transfected COS cells. This 39 kDa polypeptide lacks

CC 6 cyteine residues of the native protein yet retains similar
 CC specific enzymatic activity. The invention provides human
 CC chitinase polynucleotides and polypeptides, and methods for the
 CC recombinant production of human chitinase products, which are
 CC expected to be useful for treating fungal infections. The
 CC chitinase can be used to improve the activity of a non-chitinase
 CC antifungal agent in the treatment of candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis,
 CC mucormycosis, dermatophyoses and Pneumocystis infections.

XX Sequence 373 AA;

Query Match 61.8%; Score 34; DB 23; Length 373;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOQYTD 9
 | | | | |
 79 NFGTQKFTD 87

RESULT 25
 ABB76294
 ID ABB76294 standard; Protein; 373 AA.

XX ABB76294;

DT 12-AUG-2002 (first entry)

XX Human chitinase analogue.

XX Chitinase; enzyme; human; fungicide; antifungal; infection;
 KM candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KM paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KM chromoblastomycosis; sporotrichosis; mucormycosis; dermatophyososis;
 KM Pneumocystis; mutant; muretin.

OS Homo sapiens.
 OS Synthetic.

FH Key Location/Qualifiers
 FT Misc-difference 370

FT /note= "wild-type Pro substituted by Ser"

PN US6372212-B1.

16-APR-2002.

PF 16-JUN-1997; 97US-0877599.

PR 14-JUN-1996; 96US-0663618.

PA (ICOS-) ICOS CORP.

PI Gray PW;

DR MPI; 2002-442449/47.

XX Co-administering chitinase to improve the effectiveness of fungicidal
 XX drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 XX infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 XX Example 3; Column 43-44; 26pp; English.

XX The present sequence is the protein sequence of a human mature
 CC chitinase analogue, in which the native proline residue at amino
 CC acid position 370 (see ABB76291) is replaced by a serine residue,
 CC and in which the 72 C-terminal amino acids of the native protein
 CC are deleted. The coding region for the chitinase analogue was
 CC obtained by PCR from chitinase MO-218 cDNA (see ABUS7380), and
 CC catalytically active enzyme was produced in Escherichia coli
 CC transformants grown at 25 degrees C using plasmid pARAW0218,

CC which contains the Salmonella arabinose promoter. The invention
 CC provides human chitinase polynucleotides and polypeptides, and
 CC methods for the recombinant production of human chitinase products,
 CC which are expected to be useful for treating fungal infections. The
 CC chitinase can be used to improve the activity of a non-chitinase
 CC antifungal agent in the treatment of candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis,
 CC mucormycosis, dermatophyoses and Pneumocystis infections.

XX Sequence 373 AA;

Query Match 61.8%; Score 34; DB 23; Length 373;
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOQYTD 9
 | | | | |
 Db 79 NFGTQKFTD 87

RESULT 26
 AAW08585
 ID AAW08585 standard; Protein; 387 AA.

XX AAW08585;

DT 24-MAR-1997 (first entry)

XX Human 39 kDa chitinase.

XX Chitinase; chitotriosidase; chitin; infectious disease;
 KM gene therapy; vaccine; diagnosis; lysosomal lipidosis;
 KM Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;
 KM X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT Domain 178..198

FT /label= Catalytic_domain

PN MO9640940-A2.

PD 19-DEC-1996;

PF 06-JUN-1996; 96WO-NL00225.

PR 07-JUN-1995; 95US-0486839.

PA (UNAM) UNIV AMSTERDAM.

PI Aerts JMF;

DR MPI; 1997-118698/11.

DR N-PSDB; AAT50834.

XX New human chitinase - used to treat or prevent infection by
 PT chitin-cong. pathogens, in diagnosis and as additives to cosmetics,
 PT foods, implants etc.

PS Claim 1; Page 43-45; 58pp; English.

XX Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable
 CC to many proteases, active at pH 3-8 and up to 50 deg, and stable in
 CC the circulation. They are the result of alternative splicing of
 CC RNA, and can be produced on a large-scale in transformed host cells
 CC using cDNA clones (see also AAT50833-34) obt'd. from a human macrophage
 CC library. The chitinases, optionally expressed from a gene therapy
 CC vector, are used to treat or prevent infection by chitin-cong.
 CC pathogens (e.g. fungi, protozoa, helminths) and more generally to

CC degrade chitin. They are also useful in controlled-release drug
 CC delivery, in cosmetics, foods and dental products, for antibody
 CC prodn. and for diagnosis of diseases associated with elevated
 CC chitinase levels.

SO Sequence 387 AA;

Query Match 61.8%; Score 34; DB 18; Length 387;
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
 |||||
 Db 100 NFGTQKFTD 108

RESULT 27
 AAW08584
 ID AAW08584 standard; Protein; 466 AA.

XX AAW08584;

DE 24-MAR-1997 (first entry)

XX Human 50 kDa chitinase.

XX Chitinase; chitotriosidase; chitin; infectious disease;
 KW gene therapy; vaccine; diagnosis; lysosomal lipidosis;
 KW Gaucher's disease; leishmaniasis; sarcoidosis; multiple sclerosis;
 KW X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.
 XX Homo sapiens.

FT Key Location/Qualifiers
 FT Peptide 1..21
 FT Domain /label= Sig_peptide
 FT /label= Catalytic_domain

XX MO9640940-A2.

XX 19-DEC-1996.

XX 06-JUN-1996; 96WO-NL00225.

XX 07-JUN-1995; 95US-0466839.

XX (UNAM) UNIV AMSTERDAM.

XX Aerts JMFg;

XX WPI; 1997-118698/11.
 DR N-PSDB; AAT50833.

XX New human chitinase - used to treat or prevent infection by
 PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
 PT foods, implants etc.

XX Claim 1; Page 40-42; 56pp; English.

XX Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable
 CC to many proteases, active at pH 3-8 and up to 50 deg, and stable in
 CC the circulation. They are the result of alternative splicing of
 CC RNA, and can be produced on a large-scale in transformed host cells
 CC using cDNA clones (see also AAT50833-34) obt'd. from a human macrophage
 CC library. The chitinases, optionally expressed from a gene therapy
 CC vector, are used to treat or prevent infection by chitin-contg.
 CC pathogens (e.g. fungi, protozoa, helminths) and more generally to
 CC degrade chitin. They are also useful in controlled-release drug
 CC delivery, in cosmetics, foods and dental products, for antibody
 CC prodn. and for diagnosis of diseases associated with elevated
 CC chitinase levels.

SO Sequence 466 AA;

Query Match 61.8%; Score 34; DB 18; Length 466;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
 |||||
 Db 100 NFGTQKFTD 108

RESULT 28
 AAW40259
 ID AAW40259 standard; Protein; 466 AA.

XX AAW40259;

DE 15-JUN-1998 (first entry)

XX Human chitinase protein from clone MO-218.

XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
 KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
 KW rheumatoid arthritis; overexpression; extracellular matrix.

XX Homo sapiens.

FT Key Location/Qualifiers
 FT Peptide 1..21
 FT Protein /label= signal
 FT /label= chitinase
 FT /note= "from clone MO-218"

XX MO9747752-A1.

XX 18-DEC-1997.

XX 16-JUN-1997; 97WO-US10460.

XX 14-JUN-1996; 96US-0663618.

XX (ICOS-) ICOS CORP.

XX Gray PW;

XX WPI; 1998-052316/05.
 DR N-PSDB; AAV10435.

XX Nucleic acids encoding human chitinase - useful as antifungal
 PT agents, especially in combination with other antifungals
 PT Claim 1; Page 41-42; 63pp; English.

XX This sequence represents a novel human chitinase isolated from clone
 CC MO-218. Chitinases are useful for treating or preventing fungal infection
 CC and as immunogens for generating antibodies which are used to purify,
 CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
 CC The nucleic acid sequence of the chitinase is also useful as a probe to
 CC identify and isolate genomic DNA encoding chitinases or similar proteins,
 CC or cells expressing them or to generate transgenic ('knockout') rodents.
 CC It can also be used in hybridisation assays and to detect genetic
 CC alterations in the chitinase gene related to disease. Agents that inhibit
 CC this protein may be useful in treatment of Gaucher's disease and
 CC rheumatoid arthritis, where overexpression of the protein can damage
 CC the extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.

SO Sequence 466 AA;

Query Match 61.8%; Score 34; DB 19; Length 466;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NKGTQOYTD 9
 | |||:|
 Db 100 NFGTKKFTD 108

RESULT 29
 AA42425
 ID AA42425 standard; Protein; 466 AA.
 AC AA42425;
 DT 10-DEC-1999 (first entry)
 DE MO-218 clone of human Chitinase, amino acid sequence.
 KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
 KW organ transplant; parasite; chitin-binding; allele; vector;
 KW truncated protein; chitin binding region.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..21
 FT Peptide /label= Signal_peptide
 FT 22..445
 FT Protein /label= Mature_protein
 FT /note= "Human Chitinase"
 FT 391..445
 FT Peptide /label= Chitin_binding
 FT /note= "Chitin binding region"

PN MO9946390-A1.
 XX 16-SEP-1999.
 XX 12-MAR-1999; 99WO-US05343.
 XX 12-MAR-1999; 98US-0039198.
 XX (ICOS-) ICOS CORP.
 PA Gray PW, Tjoelker LW;
 PI MPI: 1999-551417/46.
 XX N-PSDB; AA221847.
 DR

Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals

Example 1; Page 58-59; 83pp; English.

This is the amino acid sequence of an allelic form of the human chitinase enzyme, which is capable of degrading Chitin (a linear homo polymer of beta-1,4-linked N-acetylglucosamine residues). The sequence also includes a Chitin binding peptide region, which lacks any chitinase activity, but which can be used to target anti-fungal agents to chitin containing fungal cells. Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the proteins recombinantly. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, and dermatophytoses. Chitin can be degraded by the enzyme chitinase. Use of whole chitinase

CC protein for treating infections, especially fungal infections, is CC problematic. In view of the increasing incidents of life-threatening CC fungal infection in e.g. immunocompromised individuals, there exists a CC need for identifying new compounds for treating fungal infection. The CC chitin-binding fragments of the present invention provide this need.
 XX
 SQ Sequence 466 AA;

Query Match 61.8%; Score 34; DB 20; Length 466;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
 | |||:|
 Db 100 NFGTKKFTD 108

RESULT 30
 AAE00432
 ID AAE00432 standard; Protein; 466 AA.
 AC AAE00432;
 DT 19-JUN-2001 (first entry)
 DE Human chitinase protein from clone pMO-218.
 KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-218.
 XX
 XX OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..21
 FT Peptide /label= Signal_peptide
 FT 22..466
 FT Protein /label= Human_mature_chitinase_protein
 FT 418..466
 FT Domain /label= Chitin_binding_domain
 FT 1..373
 FT Region /note= "Region with triacetylchitotriose hydrolysing activity"

PN WO200123430-A2.
 XX 05-APR-2001.
 XX 28-SEP-2000; 2000WO-US26960.
 XX 30-SEP-1999; 99US-0409918.
 XX (ICOS-) ICOS CORP.
 PA Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
 PI MPI: 2001-266141/27.
 XX N-PSDB; AAD03759.
 DR

Novel chitinase immunoglobulin fusion product, useful for treating fungal infections and reducing the amount of a non-chitinase antifungal agent needed for the treatment

Claim 1; Page 32-33; 39pp; English.

The present invention relates to a chitinase immunoglobulin (Ig) fusion product, comprising a human chitinase fused to at least a portion of an immunoglobulin chain. The fusion product is useful for treating fungal infections (mycoses) such as candidiasis, aspergillosis, blastomycosis, coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis, cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.

CC The fusion protein is useful for reducing the amount of non-chitinase
 CC antifungal agent needed to exert an antifungal activity. The fusion
 CC protein is also useful for preparing
 CC a medicament for the prophylactic or therapeutic treatment of fungal
 CC infections. Chitinase-immunoglobulin fusion product has unexpectedly
 CC improved serum half-life and formulation properties.
 CC The present sequence is human chitinase protein from clone pMO-218.
 CC Chitinase enzyme degrades chitin which is a homopolymer of
 CC beta-(1,4)-linked N-acetylglucosamine residues.

XX Sequence 466 AA;

Query Match 61.8%; Score 34; DB 22; Length 466;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTOQYTD 9
 |||::||
 Db 100 NFGTQKFTD 108

LT 31
 25903
 AAB25903 standard; Protein; 466 AA.

XX AAB25903;

DT 15-NOV-2002 (first entry)

XX Human chitinase allelic variant clone, MO-218 protein.

XX Human; chitinase; chitin-binding fragment; therapy; fungal infection;
 KW fungicide; MO-218; enzyme.

XX Homo sapiens.

OS Homo sapiens.

XX Key

XX Location/Qualifiers

XX 1..21

XX /label= Signal_peptide

XX Protein 22..466

XX /note= "Mature human MO-218 protein"

XX Domain 392..445

XX /note= "Chitin-binding domain"

XX US6399571-B1.

XX 12-MAR-1999; 99US-0267574.

XX 12-MAR-1998; 98US-0039198.

XX (ICOS-) ICOS CORP.

XX Gray PW, Tjoelker LW;

XX WPI; 2002-606631/65.

XX N-PSDB; AAD42616.

XX New chitinase derived polypeptides used for treating fungal infections
 PT and for detecting and binding chitin -

XX Claim 1; Column 35-36; 32pp; English.

XX The invention relates to chitinase derived polypeptide, chitin-binding
 CC fragments of human chitinase, fragment analogues, purified and isolated
 CC polynucleotides encoding such fragments. The invention is used for
 CC treating fungal infections and for detecting and binding chitin. The
 CC present sequence is human chitinase allelic variant clone, MO-218
 CC protein.

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

XX Sequence 466 AA;

Query Match 61.8%; Score 34; DB 23; Length 466;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTOQYTD 9
 |||::||
 Db 100 NFGTQKFTD 108

RESULT 32

ID ABB76291 standard; Protein; 466 AA.

XX ABB76291;

DT 12-AUG-2002 (first entry)

XX Human chitinase.

XX Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis.

XX Homo sapiens.

OS Homo sapiens.

XX Key

XX Location/Qualifiers

XX 1..21

XX /label= Signal_peptide

XX Protein 22..466

XX /label= Mature_protein

XX US6372212-B1.

XX 16-APR-2002.

XX 16-JUN-1997; 97US-0877599.

XX 14-JUN-1996; 96US-0663618.

XX (ICOS-) ICOS CORP.

XX Gray PW;

XX WPI; 2002-442449/47.

XX N-PSDB; ABL57380.

XX Co-administering chitinase to improve the effectiveness of fungicidal
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -

XX Claim 1; Column 27-30; 26pp; English.

XX The present sequence is the protein sequence of human chitinase
 CC as predicted from isolated cDNA clone MO-218 (see ABL57380). It
 CC differs in only 1 amino acid residue from the chitinase sequence
 CC (see ABB76292) deduced from a second cDNA clone, having glycine
 CC at position 81 of the mature protein. Northern blots showed
 CC highest chitinase gene expression in lung and ovary tissues.

XX Expression in lung is consistent with a protective role against
 CC pathogenic organisms that contain chitin. The invention provides
 CC human chitinase polynucleotides and polypeptides, and materials and
 CC methods for the recombinant production of human chitinase products,
 CC which are expected to be useful as products for treating fungal
 CC infections or for the development of such products. Human
 CC chitinase has a synergistic effect on the actions of other
 CC fungicides. It can be administered to improve the antifungal
 CC activity of a non-chitinase antifungal agent, especially
 CC amphotericin B or itraconazole, in the treatment of a fungal
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,

CC dermatophytes and Pneumocystis infections (all claimed). In
CC particular, the fungal infection involves Candida, Aspergillus
CC and/or Cryptococcus spp., whose growth is not effectively
CC inhibited by contact with human chitinase alone.

XX SQ Sequence 466 AA;

Query Match 61.8%; Score 34; DB 23; Length 466;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOOYTD 9
DB 100 NKGTOOYTD 108

RESULT 33

ABR92136
ABR92136 standard; Protein; 466 AA.

ABR92136;

XX DT 31-MAY-2002 (first entry)

XX DE Herbicidally active polypeptide SEQ ID NO 1347.

XX KW Herbicidal; plant; agriculture; herbicide.

XX OS Arabidopsis thaliana.

XX PN WO200210210-A2.

XX PD 07-FEB-2002.

XX PF 28-AUG-2001; 2001WO-BP09892.

XX PR 28-AUG-2001; 2001WO-BP09892.

XX PA (FARB) BAYER AG.

XX PI Tietjen K, Weidner M;

XX DR WPI; 2002-269010/31.

XX PT Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -

PS Claim 5; SEQ ID NO 1347; 261bp + Sequence Listing; English.

XX The invention relates to identifying target proteins
XX (ABR90790-ABR94016) for herbicidally active compounds, comprising
XX aligning and comparing nucleic acid or amino acid sequences from plant
XX with nucleic acid or amino acid sequences from non-plant organisms using
XX suitable search parameters, where plant sequences having an E-value
XX greater by a factor of 3 than the E-value of most similar non-plant
XX sequences are selected. The polypeptides or nucleic acids encoding them
XX are useful for identifying modulators. The identified modulators are
XX useful as herbicides.

XX SQ Sequence 466 AA;

Query Match 61.8%; Score 34; DB 23; Length 466;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KGTQOYTD 10
DB 86 KGAQLYTD 94

RESULT 34

AAV23206
ID AAV23206 standard; Protein; 493 AA.

XX AC AAV23206;

XX DT 24-AUG-1999 (first entry)

XX DE Amino acid sequence of Cry3Bb.60 polypeptide.

XX KW Cry3Bb; mutant; insecticidal activity; insecticidal specificity;

XX KW coleoptera; southern corn rootworm; western corn root worm;

XX KW Diabrotica undecimpunctata howardi Barber; transgenic plant;

XX KW Diabrotica virgifera vergifera LeConte; insecticide resistance.

XX OS Synthetic.

XX OS Bacillus thuringiensis.

XX PN WO9931248-A1.

XX PD 24-JUN-1999.

XX PF 17-DEC-1998; 98WO-US26852.

XX PR 18-DEC-1997; 97US-0996441.

XX PR 18-DEC-1997; 97US-0993170.

XX PR 18-DEC-1997; 97US-0993172.

XX PR 18-DEC-1997; 97US-0993175.

XX PA (ECOG-) ECOGEN INC.

XX PA (MONS) MONSANTO CO.

XX PI Brussels SW, Bryson JW, English L, Kulesza CA, Malvar TW;

XX PI Romano C, Slatin SL, Von Tersch MA, Walters FS;

XX DR WPI; 1999-395184/33.

XX PT Insecticidal Bacillus thuringiensis proteins

XX PS Claim 39; Page 462-464; 512pp; English.

XX AAV23172-Y23206, and AAV23208-X23209 represent new Bacillus
XX thuringiensis Cry3Bb mutant proteins. The specification also describes
XX methods of altering Bacillus thuringiensis Cry3Bb. The B. thuringiensis
XX Cry3Bb polypeptide was modified to have improved insecticidal activity
XX or enhanced insecticidal specificity against a target insect. The
XX modification comprises at least one amino acid substitution, addition,
XX or deletion in the primary sequence of the native or unmodified Cry3Bb
XX polypeptide, wherein the substitution or deletion occurs at a position
XX corresponding to from about amino acids 1-365 of the unmodified
XX polypeptide sequence (AAV23207 represents the wild type Cry3Bb
XX protein). The polypeptide can be used to kill coleopteran pests,
XX especially by application to the environment. It is especially
XX useful against southern corn rootworm and western corn root worm,
XX (Diabrotica undecimpunctata howardi Barber, and Diabrotica virgifera
XX vergifera LeConte respectively). The mutant Cry3Bb polynucleotides
XX can also be used to produce transgenic plants with increased
XX insecticide resistance.

XX SQ Sequence 493 AA;

Query Match 61.8%; Score 34; DB 20; Length 493;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 76 KLTQOYTD 83

RESULT 35

AAK26574
ID AAK26574 standard; Protein; 572 AA.

XX

AC	AAK26574;
XX	
DT	25-MAR-2003 (updated)
DT	28-JAN-1993 (first entry)
XX	
DE	Sequence encoded by CryIIIC(b) gene.
XX	
KM	Insecticidal crystal protein; coleoptera; Colorado potato beetle;
XX	Dibrotica; 70kd insecticidal protein.
XX	
OS	Bacillus thuringiensis strain EG5144.
PN	M09213954-A1.
PD	
FP	20-AUG-1992.
PP	03-JAN-1992; 92WO-US00040.
PR	31-JAN-1991; 91US-064956Z.
PP	23-DEC-1991; 91US-081359Z.
XX	(ECOG-) ECOGEN INC.
PI	Donovan WP, Ruper MJ, Slaney AC;
DR	WPI, 1992-300040/36.
XX	N-PSDB; AAQ27945.
PT	Bacillus thuringiensis conty. CryIIIC (B) gene in plasmid - used to make insecticide against coleopterian insects
PS	Claim 4; Fig 1; 91pp; English.
CC	A comparison of the nucleotide base sequence of the CryIIIC(b) gene coding region with the corresp. coding region of the prior art CryIIIA gene indicates significant differences between the two genes. The CryIIIC(b) gene is only 76% homologous (positionally identical) with the CryIIIA gene. The sequence of the CryIIIC(b) gene coding region with the corresp. coding region of the CryIIIB gene obtained from recently discovered B.t. strain EG2838 (NRRL Accession No. B-18603) indicates that the CryIIIC(b) gene is 96% homologous (positionally identical) with the CryIIIB gene. (Updated on 25-MAR-2003 to correct FN field.)
SQ	Sequence 572 AA;
Query Match	61.8%; Score 34; DB 13; Length 572;
Best Local Similarity	87.5%; Pred. NO. 4.4e+02;
Matches 7; Conservative	0; Mismatches 1; Indels 0; Gaps 0
CY	2 KGTQQYT D 9
Db	195 KLTQQYT D 202
RESULT 36	
ID	AAK06460
AC	AAK06460 standard; protein; 651 AA.
XX	
AC	AAK06460;
XX	
DT	25-MAR-2003 (updated)
DT	03-JAN-1991 (first entry)
XX	
DE	BtPGS1208 protoxin.
XX	
KM	Toxin; crystal; insecticide.
XX	
OS	Bacillus thuringiensis strain PGS1208.
XX	
FT	Key Location/Qualifiers
FT	Protein 58..651
FT	/label=BtPGS1208

XX	EPJ82990-A.
PN	
XX	22-AUG-1990.
PD	
XX	15-FEB-1989; 89EP-0400428.
PF	
XX	15-FEB-1989; 89EP-0400428.
PR	
XX	(PLBZ) PLANT GENETIC SYSTEMS NV.
PA	
XX	Peferoen M, Lambert B, Joos H;
PI	
XX	WPI, 1990-255362/34.
DR	N-PSDB; AAQ05679.
XX	New Bacillus thuringiensis strains - producing toxin active against Coleoptera
PT	
XX	Claim 2, Fig 1, 30pp; English.
PS	
CC	The BtPGS1208 strain was isolated from grain dust and was deposited at the DSM (19/1/1989) under Acc. No. 5131. The protoxin (74 kD) can be digested with trypsin to produce the mature protein (67 kD). The proteins formed during sporulation are packaged in flat rhomboid crystals. Plants contg. the gene encoding the toxin are resistant to Coleoptera pests while the crystals are useful as insecticides. The gene can also be truncated and used to produce chimaeric genes together with the btPGS1245 gene (AAQ05680).
CC	See also AHR06461.
CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 651 AA;
Query Match	61.8%; Score 34; DB 11; Length 651;
Best Local Similarity	87.5%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	2 KGTQQYT D 9 Db 235 KLRTQQYT D 242
RESULT 37	
ID	AAR33769 AAR33769 standard; Protein; 651 AA.
XX	
AC	AAR33769;
XX	
DT	25-MAR-2003 (updated)
DT	15-JUL-1993 (first entry)
XX	
DE	Bt isolate 43F.
XX	
KW	Lepidopteran-active; toxin; T1; R1; plant; cell; Bt.
XX	
OS	Bacillus thuringiensis.
PN	
WO	WO9304587-A1.
XX	
PD	18-MAR-1993.
XX	
PF	11-SEP-1992; 92WO-US07697.
PR	12-SEP-1991; 91US-0758020.
PR	08-SEP-1992; 92US-0941650.
XX	
PA	(MYCO) MYCOGEN CORP.
XX	
PI	Bradfish GA, Uyeda KA;
XX	
WPI	WPI, 1993-100566/12.
DR	N-PSDB; AAQ38654.
OR	

XX Controlling lepidopteran pests - using compsn. of *Bacillus*
 PT *thuringiensis* strains or plants or microorganisms transformed
 PT with their toxin genes
 PS Claim 13; Page 25-27; 38pp; English.
 XX The sequences given in AAM06419 represent lepidopteran-active toxins.
 CC The DNA encoding these sequences were used within a Ti or Ri plasmid,
 CC to transform plant cells. Whole plants can then be regenerated from
 CC the transformed cells. The toxin may also be produced by cloning
 CC *Bacillus thuringiensis* (Bt). It may then be applied directly to
 CC the plant locus.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC
 SQ Sequence 651 AA;
 QY Query Match 61.8%; Score 34; DB 14; Length 651;
 DB Best Local Similarity 87.5%; Pred. No. 5.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 2 KGTQGYTD 9
 235 KLTQGYTD 242

RESULT 38
 AAM06419
 ID AAM06419 standard; Protein; 651 AA.
 AC AAM06419;
 XX
 XX 28-JAN-1997 (first entry)
 DT
 XX Anti-scarab pest toxin 43F.
 DE
 XX Toxin 50C(a); scarab pest; toxin; insect; scarabaeidae; pest control;
 KW larval stage insect; grain; tuberos crop; white grub; chafer grub;
 KW cyclocephala; popillia; 43F gene.
 KM
 OS *Bacillus thuringiensis* strain tolworthi.
 OS
 XX US5554534-A.
 PN
 XX 10-SEP-1996.
 PD
 XX 30-SEP-1994; 94US-0315468.
 XX
 XX 01-FEB-1993; 93US-0014941.
 PR 16-DEC-1991; 91US-0808316.
 PR 30-JAN-1993; 92US-0828430.
 PR 30-SEP-1994; 94US-0315468.
 XX
 XX (MYCO) MYCOGEN CORP.
 PA
 XX Fonceerrada L, Michaels TE, Narva KE;
 PI
 XX WPI; 1996-424659/42.
 DR N-PSDB; AAT43223.
 DR
 XX New nucleic acid encoding *B. thuringiensis* toxin active against
 PT scarab(a) - also related toxin and transformed microbes, effective
 PT against adult pests and their larvae
 PS Example 3; Column 37-42; 24pp; English.
 XX AAM06417-M06419 represent toxins that are active against scarab pests.
 CC The DNA encoding this sequence was isolated from the *Bacillus*
 CC *thuringiensis* strain tolworthi. Insects in the family Scarabaeidae
 CC constitute a serious pest control problem, especially when destructive
 CC larval stage insects infest high value turf found in golf courses,
 CC playing fields and lawns. The larvae of many species also attack grains,

CC tuberos crops, and ornamentals. The larvae are known as white grubs, or
 CC chafer grubs, and can be found in decaying organic matter, or in the soil
 CC where they consume plant roots. In Europe and the U.S. populations of
 CC these larvae and adults have developed resistance to chemical
 CC insecticides such as the organochlorines and DDT. These toxin sequences,
 CC and intact cells that are capable of expressing these proteins, can be
 CC used to control many pests of the family scarabaeidae, such as species of
 CC *Cyclocephala*, and *Popillia*. The toxins are active against larvae
 CC (present in soil) and against adults.
 CC
 SQ Sequence 651 AA;
 QY Query Match 61.8%; Score 34; DB 17; Length 651;
 DB Best Local Similarity 87.5%; Pred. No. 5.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 2 KGTQGYTD 9
 235 KLTQGYTD 242

RESULT 39
 AAY23197
 ID AAY23197 standard; Protein; 651 AA.
 AC AAY23197;
 XX
 XX 24-AUG-1999 (first entry)
 DT
 XX Amino acid sequence of Cry3Bb.11048 polypeptide.
 DE
 XX Cry3Bb; mutant; insecticidal activity; insecticidal specificity;
 KW coleoptera; southern corn rootworm; western corn root worm;
 KW diptera; undecimnotata howard; barber; transgenic plant;
 KW diptera; virgiferla verigifera leconte; insecticide resistance.
 KM
 OS Synthetic.
 OS *Bacillus thuringiensis*.
 OS
 XX WO9931248-A1.
 PN
 XX 24-JUN-1999.
 PD
 XX 17-DEC-1998; 98WO-US26852.
 PF
 XX 18-DEC-1997; 97US-0996441.
 PR 18-DEC-1997; 97US-0993170.
 PR 18-DEC-1997; 97US-0993722.
 PR 18-DEC-1997; 97US-0993775.
 XX
 XX (ECOG) ECOGEN INC.
 PA (MONS) MONSANTO CO.
 PI
 XX Brunsack SM, Bryson JW, English L, Kulesza CA, Malvar TW;
 PI Romano C, Slatin SL, Von Tersch MA, Walters FS;
 XX
 XX WPI; 1999-395184/33.
 DR
 XX Insecticidal *Bacillus thuringiensis* proteins
 PI
 XX Claim 39; Page 412-415; 512pp; English.
 PS
 XX AAY23197-23206, and AAY23208-X23209 represent new *Bacillus*
 CC *thuringiensis* Cry3Bb mutant proteins. The specification also describes
 CC methods of altering *Bacillus thuringiensis* Cry3Bb. The *B. thuringiensis*
 CC Cry3Bb polypeptide was modified to have improved insecticidal activity
 CC or enhanced insecticidal specificity against a target insect. The
 CC modification comprises at least one amino acid substitution, addition,
 CC or deletion in the primary sequence of the native or unmodified Cry3Bb
 CC polypeptide, wherein the substitution or deletion occurs at a position
 CC corresponding to from about amino acids 1-365 of the unmodified
 CC polypeptide sequence (AAY23207 represents the wild type Cry3Bb
 CC protein). The polypeptide can be used to kill coleopteran pests,

CC especially by application to the environment. It is especially
 CC useful against southern corn rootworm and western corn root worm,
 CC (*Diabrotica undecimpunctata howardi* Barber, and *Diabrotica virgifera*
 CC *virgifera* LeConte respectively). The mutant cry3Bb polynucleotides
 CC can also be used to produce transgenic plants with increased
 CC insecticide resistance.

XX Sequence 651 AA;

Query Match 61.8%; Score 34; DB 20; Length 651;

Best Local Similarity 87.5%; Pred. No. 5.1e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
 |||||
 DB 234 KLTQOYTD 241

RESULT 40

AA23199 AAY23199 standard; Protein; 651 AA.

AA23199;

24-AUG-1999 (first entry)

Amino acid sequence of Cry3Bb.11057 polypeptide.

Cry3Bb; mutant; insecticidal activity; insecticidal specificity;

coleoptera; southern corn rootworm; western corn root worm;

Diabrotica undecimpunctata howardi Barber; transgenic plant;

Diabrotica virgifera LeConte; insecticide resistance.

Synthetic.

Bacillus thuringiensis.

MO9931248-A1.

24-JUN-1999.

17-DEC-1998; 98WO-US26852.

18-DEC-1997; 97US-0996441.

18-DEC-1997; 97US-0993170.

18-DEC-1997; 97US-0993722.

18-DEC-1997; 97US-0993775.

(ECOG-) ECOGEN INC.

(MONS) MONSANTO CO.

Brussock SM, Bryson JW, English L, Kulesza CA, Malvar TM;

Romano C, Slatin SL, Von Tersch MA, Walters FS;

WPI; 1999-395184/33.

Insecticidal *Bacillus thuringiensis* proteins

Claim 39; Page 424-426; 512pp; English.

CC *virgifera* LeConte respectively). The mutant cry3Bb polynucleotides
 CC can also be used to produce transgenic plants with increased
 CC insecticide resistance.

XX Sequence 651 AA;

Query Match 61.8%; Score 34; DB 20; Length 651;

Best Local Similarity 87.5%; Pred. No. 5.1e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
 |||||
 DB 234 KLTQOYTD 241

Search completed: August 20, 2003, 09:29:36
 Job time : 35.0206 secs

THIS PAGE BLANK (USE TOP)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:07 ; Search time 11.0309 Seconds
(without alignments)
38.357 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKG2QYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

1 number of hits satisfying chosen parameters: 328717

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	395	3	US-08-856-841-13
2	55	100.0	398	3	US-08-856-841-21
3	55	100.0	415	3	US-08-856-841-20
4	55	100.0	543	3	US-08-856-841-22
5	55	100.0	756	4	US-09-438-268-4
6	35	63.6	61	4	US-09-328-352-7978
7	35	63.6	415	4	US-09-554-999-2
8	34	61.8	187	4	US-09-095-855-203
9	34	61.8	187	4	US-09-205-426-203
10	34	61.8	373	3	US-09-039-198A-14
11	34	61.8	373	3	US-09-039-198A-15
12	34	61.8	373	4	US-08-877-589-14
13	34	61.8	373	4	US-08-877-589-15
14	34	61.8	373	4	US-09-267-574-14
15	34	61.8	373	4	US-09-267-574-15
16	34	61.8	387	2	US-08-486-839-6
17	34	61.8	387	2	US-08-151-011-6
18	34	61.8	387	4	US-09-343-623-6
19	34	61.8	466	2	US-08-486-839-4
20	34	61.8	466	3	US-09-151-011-4
21	34	61.8	466	3	US-09-039-198A-2
22	34	61.8	466	4	US-09-343-623-4
23	34	61.8	466	4	US-08-877-589-2
24	34	61.8	466	4	US-09-267-574-2
25	34	61.8	493	4	US-08-996-441B-70
26	34	61.8	493	3	US-08-993-722A-70
27	34	61.8	493	3	US-08-993-170A-70

28	34	61.8	493	3	US-08-993-775B-70	Sequence 70, Appl
29	34	61.8	651	1	US-08-315-468-6	Sequence 6, Appl
30	34	61.8	651	1	US-08-996-441B-52	Sequence 52, Appl
31	34	61.8	651	3	US-08-996-441B-56	Sequence 56, Appl
32	34	61.8	651	3	US-08-996-441B-58	Sequence 58, Appl
33	34	61.8	651	3	US-08-993-722A-52	Sequence 52, Appl
34	34	61.8	651	3	US-08-993-722A-56	Sequence 56, Appl
35	34	61.8	651	3	US-08-993-722A-58	Sequence 58, Appl
36	34	61.8	651	3	US-08-993-170A-52	Sequence 52, Appl
37	34	61.8	651	3	US-08-993-170A-56	Sequence 56, Appl
38	34	61.8	651	3	US-08-993-170A-58	Sequence 58, Appl
39	34	61.8	651	3	US-08-993-775B-52	Sequence 52, Appl
40	34	61.8	651	3	US-08-993-775B-56	Sequence 56, Appl
41	34	61.8	651	3	US-08-993-775B-58	Sequence 58, Appl
42	34	61.8	651	3	US-07-941-650A-4	Sequence 4, Appl
43	34	61.8	652	3	US-08-996-441B-2	Sequence 2, Appl
44	34	61.8	652	3	US-08-996-441B-4	Sequence 4, Appl
45	34	61.8	652	3	US-08-996-441B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-856-841-13
Sequence 13, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 395
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE

HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERTHEMA INFECTION)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
OTHER INFORMATION: mass spectrometry
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 13:
US-08-856-841-13
Query Match 100.0%; Score 55; DB 3; Length 395;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTOOYTDQ 10
Db 256 NKGTOOYTDQ 265
RESULT 2
US-08-856-841-21
Sequence 21, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856, 841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214, 658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917, 096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 398
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 21:
US-08-856-841-21

Query Match 100.0%; Score 55; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
|||||
245 NKGTOOYTDQ 254

RESULT 3

US-08-856-841-20
; Sequence 20, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y. E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 20:
US-08-856-841-20

Query Match 100.0%; Score 55; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOOYTDQ 10
|||||
256 NKGTOOYTDQ 265

RESULT 4

US-08-856-841-22
; Sequence 22, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543
; TYPE: AMINO ACID

TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHEICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 100.0%; Score 55; DB 3; Length 543;
Best Local Similarity 100.0%; Pred. No. 0.0057;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQ 10
DB 404 NKGTQOYTDQ 413

RESULT 5
US-09-438-268-4
Sequence 4, Application US/09438268
Patent No. 6491907
GENERAL INFORMATION:
APPLICANT: Rabinowitz, Joseph E.
APPLICANT: Samulski, Richard J.
TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
FILE REFERENCE: 5470-186
CURRENT APPLICATION NUMBER: US/09/438,268
CURRENT FILING DATE: 1999-11-10
EARLIER APPLICATION NUMBER: 60/107,840
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/123,651
NUMBER OF SEQ ID NOS: 59
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 4
LENGTH: 756
TYPE: PRT
ORGANISM: Virus
US-09-438-268-4

Query Match 100.0%; Score 55; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQ 10
DB 1111111111

DB 597 NKGTQOYTDQ 606
RESULT 6
US-09-328-352-7978
Sequence 7978, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7978
LENGTH: 61
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7978

Query Match 63.6%; Score 35; DB 4; Length 61;
Best Local Similarity 70.0%; Pred. No. 5.2;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQ 10
DB 5 NKGTORSQDQ 14

RESULT 7
US-09-554-999-2
Sequence 2, Application US/09554999
Patent No. 6465186
GENERAL INFORMATION:
APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases from Gram Positive Organisms
FILE REFERENCE: GC389-US
CURRENT APPLICATION NUMBER: US/09/554,999
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: PCT/US98/27040
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: GB 9727471.6
PRIOR FILING DATE: 1997-12-30
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-554-999-2

Query Match 63.6%; Score 35; DB 4; Length 415;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 NKGTQOYTDQ 10
DB 208 NQKKPYTDQ 217

RESULT 8
US-09-095-855-203
Sequence 203, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
NUMBER OF SEQUENCES: Treatment and Diagnosis of Mycobacterial Infections
208

CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000,1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 203:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-855-203

Query Match 61.8%; Score 34; DB 3; Length 187;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
DB 161 DKSTHOYTNO 170

RESULT 9
US-09-205-426-203
Sequence 203, Application US/09205426
Patent No. 6406704
GENERAL INFORMATION:
APPLICANT: Watson, James D.
TITLE OF INVENTION: Compounds and Methods for Treatment and
FILE REFERENCE: 11000,1002c4
CURRENT APPLICATION NUMBER: US/09/205,426
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: 09/095,855
EARLIER FILING DATE: 1998-06-11
EARLIER APPLICATION NUMBER: 08/997,362
EARLIER FILING DATE: 1997-12-23
EARLIER APPLICATION NUMBER: 08/873,970
EARLIER FILING DATE: 1997-06-12
EARLIER APPLICATION NUMBER: 08/705,347
EARLIER FILING DATE: 1996-08-29
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187

TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
US-09-205-426-203

Query Match 61.8%; Score 34; DB 4; Length 187;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
DB 161 DKSTHOYTNO 170

RESULT 10
US-09-039-198A-14
Sequence 14, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-14

Query Match 61.8%; Score 34; DB 3; Length 373;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTOQYTD 9
DB 79 NFGTOKFTD 87

RESULT 11
US-09-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
ADDRESSEE: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 61.8%; Score 34; DB 3; Length 373;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOYTD 9
| | | | |
Db 79 NFGTKFTD 87

RESULT 12
US-08-877-599-14
Sequence 14, Application US/0887599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-877-599-14

Query Match 61.8%; Score 34; DB 4; Length 373;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOYTD 9
| | | | |
Db 79 NFGTKFTD 87

RESULT 13
US-08-877-599-15
Sequence 15, Application US/0887599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-877-599-15

Query Match 61.8%; Score 34; DB 4; Length 373;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOYTD 9
| | | | |

Db 79 NFGTOKFTD 87

RESULT 14

US-09-267-574-14
Sequence 14, Application US/09267574
Patent No. 6399571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITTIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: 09/039,198
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-09-267-574-14

Query Match

61.8%; Score 34; DB 4; Length 373;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;Qy 1 NKGTOQYTD 9
Db 79 NFGTOKFTD 87

RESULT 15

US-09-267-574-15
Sequence 15, Application US/09267574
Patent No. 6399571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITTIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/09/267,574
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: 09/039,198
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-09-267-574-15Qy 1 NKGTOQYTD 9
Db 79 NFGTOKFTD 87Query Match 61.8%; Score 34; DB 4; Length 373;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;Qy 1 NKGTOQYTD 9
Db 79 NFGTOKFTD 87

RESULT 16

US-08-486-839-6
Sequence 6, Application US/08486839
Patent No. 5928928

GENERAL INFORMATION:

APPLICANT: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

SOFTWARE:

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07 - June - 1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-486-839-6Query Match 61.8%; Score 34; DB 2; Length 387;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;Qy 1 NKGTOQYTD 9
Db 100 NFGTOKFTD 108

RESULT 17

US-09-151-011-6
Sequence 6, Application US/09151011
Patent No. 6057142

GENERAL INFORMATION:

TITLE OF INVENTION: A Human Chitinase, Its Recombinant
production, Its Use For Decomposing Chitin, Its Use in
Therapy or Prophylaxis Against Infection Diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998
ATTORNEY/AGENT INFORMATION:
NAME: Morris, Robert C.
REGISTRATION NUMBER: 42,910
REFERENCE/DOCKET NUMBER: 294-32 DIV
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-151-011-6

Query Match 61.8%; Score 34; DB 3; Length 387;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 NKGTOOYTD 9
100 NFGTOKFTD 108

LT 18
US-09-343-623-6
Sequence 6, Application US/09343623
Patent No. 6303118
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.
TITLE OF INVENTION: 16
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-343-623-6

Query Match 61.8%; Score 34; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 NKGTOOYTD 9
100 NFGTOKFTD 108

DB 100 NFGTOKFTD 108

RESULT 19
US-08-486-839-4
Sequence 4, Application US/08486839
Patent No. 5928928
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.
TITLE OF INVENTION: 16
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07 - June - 1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-486-839-4

Query Match 61.8%; Score 34; DB 2; Length 466;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 NKGTOOYTD 9
100 NFGTOKFTD 108

RESULT 20
US-09-151-011-4
Sequence 4, Application US/09151011
Patent No. 6057142
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A Human Chitinase, Its Recombinant
production, Its Use For Decomposing Chitin, Its Use in
Therapy or Prophylaxis Against Infection Diseases.
TITLE OF INVENTION: 16
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998
ATTORNEY/AGENT INFORMATION:
NAME: Morris, Robert C.
REGISTRATION NUMBER: 42,910
REFERENCE/DOCKET NUMBER: 294-32 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-151-011-4

Query Match 61.8%; Score 34; DB 3; Length 466;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NKGTQGYTD 9
||||:|
Db 100 NFGTKFTD 108

RESULT 21
US-09-039-198A-2
Sequence 2, Application US/09039198A
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, LA-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-039-198A-2

Query Match 61.8%; Score 34; DB 3; Length 466;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NKGTQGYTD 9
||||:|
Db 100 NFGTKFTD 108

RESULT 22
US-09-343-623-4
Sequence 4, Application US/09343623
Patent No. 6303118
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-343-623-4

Query Match 61.8%; Score 34; DB 4; Length 466;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NKGTQGYTD 9
||||:|
Db 100 NFGTKFTD 108

RESULT 23
US-08-877-599-2
Sequence 2, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/877,599

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/663,618

FILING DATE: 14-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 27866/33994

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 466 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-877-599-2

Query Match

Best Local Similarity 61.8%; Score 34; DB 4; Length 466;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

Qy 1 NKGTOOYTD 9

Db 100 NFGTOKFTD 108

RESULT 25
US-08-996-441B-70
Sequence 70, Application US/08996441B

Patent No. 6023013

GENERAL INFORMATION:

APPLICANT: English, Leigh H.

APPLICANT: Brusock, Susan M.

APPLICANT: Malvar, Thomas M.

APPLICANT: Bryson, James W.

APPLICANT: Kulesza, Caroline A.

APPLICANT: Walters, Frederick S.

APPLICANT: Slatin, Stephen L.

APPLICANT: Von Terach, Michael A.

APPLICANT: Romano, Charles

TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSER: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996,441B

FILING DATE: 18-DEC-1997

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Kitchell, Barbara S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: MECO:151

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/418-3000

TELEFAX: 512/474-7577

INFORMATION FOR SEQ. ID NO. 70:

SEQUENCE CHARACTERISTICS:

LENGTH: 493 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-996-441B-70

Query Match

Best Local Similarity 61.8%; Score 34; DB 3; Length 493;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

Db 76 KLTQOYTD 83

Qy 2 KGTQOYTD 9

NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,722A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-722A-70

Query Match 61.8%; Score 34; DB 3; Length 493;
Best Local Similarity 87.5%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 76 KLTQOYTD 83

RESULT 27
US-08-993-170A-70
Sequence 70, Application US/08993170A
Patent No. 6063597
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-170A-70

Query Match 61.8%; Score 34; DB 3; Length 493;
Best Local Similarity 87.5%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 76 KLTQOYTD 83

RESULT 28
US-08-993-775B-70
Sequence 70, Application US/08993775B
Patent No. 6077824
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
DELTA-ENDOTOXINS AGAINST INSECT PESTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-70

Query Match 61.8%; Score 34; DB 3; Length 493;
Best Local Similarity 87.5%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
| | | | |
Db 76 KLTQOYTD 83

RESULT 29
US-08-315-468-6
; Sequence 6, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaelis, Tracy Ellis
; APPLICANT: Foncerea, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: tolworthi
; IMMEDIATE SOURCE: 43P
; CLONE: E. coli XL1-Blue (pm1,98-4), NRRL B-18291
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..651
US-08-315-468-6
Query Match 61.8%; Score 34; DB 1; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
| | | | |
Db 235 KLTQOYTD 242

RESULT 30
US-08-996-441B-52
; Sequence 52, Application US/08996441B
; Patent No. 6023013
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brusock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Terreb, Michael A.
; TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,441B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO.151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-996-441B-52
Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
| | | | |
Db 234 KLTQOYTD 241
RESULT 31
US-08-996-441B-56
; Sequence 56, Application US/08996441B
; Patent No. 6023013
; GENERAL INFORMATION:
; APPLICANT: English, Leigh H.
; APPLICANT: Brusock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.

APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-56

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 234 KLTQOYTD 241

RESULT 32
US-08-996-441B-58
Sequence 58, Application US/08996441B
Patent No. 6023013
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
TITLE OF INVENTION: INSECT-RESISTANT TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996,441B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:151
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-441B-58

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 234 KLTQOYTD 241

RESULT 33
US-08-993-722A-52
Sequence 52, Application US/08993722A
Patent No. 6060594
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Tersch, Michael A.
APPLICANT: Romano, Charles
TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,722A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3106
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-993-722A-52

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
 | | | | |
DB 234 KLTQOYTD 241

RESULT 34

US-08-993-722A-56

; Sequence 56, Application US/08993722A
; Patent No. 6060594

; GENERAL INFORMATION:

; APPLICANT: English, Leigh H.
; APPLICANT: Brusock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Romano, Charles
; APPLICANT: Von Teresch, Michael A.
; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
; TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,722A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3106
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-722A-56

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
 | | | | |
DB 234 KLTQOYTD 241

RESULT 35

US-08-993-722A-58

; Sequence 58, Application US/08993722A
; Patent No. 6060594

; GENERAL INFORMATION:

; APPLICANT: English, Leigh H.
; APPLICANT: Brusock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Romano, Charles
; APPLICANT: Von Teresch, Michael A.
; TITLE OF INVENTION: NUCLEIC ACID SEGMENTS ENCODING MODIFIED
; TITLE OF INVENTION: COLEOPTERAN-TOXIC CRYSTAL PROTEINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,722A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: MECO:149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3106
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-722A-58

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
 | | | | |
DB 234 KLTQOYTD 241

RESULT 36

US-08-993-170A-52

; Sequence 52, Application US/08993170A
; Patent No. 6063597

; GENERAL INFORMATION:

; APPLICANT: English, Leigh H.
; APPLICANT: Brusock, Susan M.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Bryson, James W.
; APPLICANT: Kulesza, Caroline A.
; APPLICANT: Walters, Frederick S.
; APPLICANT: Slatin, Stephen L.
; APPLICANT: Von Teresch, Michael A.
; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
; TITLE OF INVENTION: COLEOPTERAN INSECTS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas

COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993.170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-170A-52

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 234 KLTQOYTD 241

RESULT 37
US-08-993-170A-56
Sequence 56, Application US/08993170A
Patent No. 6063597
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Teresch, Michael A.
TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993.170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000

TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-170A-56

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 234 KLTQOYTD 241

RESULT 38
US-08-993-170A-58
Sequence 58, Application US/08993170A
Patent No. 6063597
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Teresch, Michael A.
TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993.170A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-170A-58

Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 234 KLTQOYTD 241

RESULT 39
US-08-993-775B-52
Sequence 52, Application US/08993775B
Patent No. 6077824
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Teresch, Michael A.
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
DELTA-ENDOTOXINS AGAINST INSECT PESTS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-52
Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
DB 234 KLTQOYTD 241

RESULT 40
US-08-993-775B-56
Sequence 56, Application US/08993775B
Patent No. 6077824
GENERAL INFORMATION:
APPLICANT: English, Leigh H.
APPLICANT: Brussock, Susan M.
APPLICANT: Malvar, Thomas M.
APPLICANT: Bryson, James W.
APPLICANT: Kulesza, Caroline A.
APPLICANT: Walters, Frederick S.
APPLICANT: Slatin, Stephen L.
APPLICANT: Von Teresch, Michael A.
TITLE OF INVENTION: METHODS FOR IMPROVING THE ACTIVITY OF
DELTA-ENDOTOXINS AGAINST INSECT PESTS
NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,775B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:150
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-775B-56
Query Match 61.8%; Score 34; DB 3; Length 651;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
DB 234 KLTQOYTD 241

Search completed: August 20, 2003, 09:34:28
Job time : 12.0309 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:30:58 ; Search time 40.6186 Seconds
(without alignments)
32.488 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKG2QYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 497079

Maximum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCVS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	756	US-10-205-942-4	Sequence 4, Appli
2	35	63.6	233	US-10-188-586A-7	Sequence 7, Appli
3	35	63.6	415	US-10-188-586A-2	Sequence 2, Appli
4	34	61.8	187	US-10-205-979-49	Sequence 49, Appli
5	34	61.8	187	US-10-051-643-203	Sequence 203, Appli
6	34	61.8	373	US-10-161-547-14	Sequence 14, Appli
7	34	61.8	373	US-10-161-547-15	Sequence 15, Appli
8	34	61.8	445	US-10-004-219B-10	Sequence 10, Appli
9	34	61.8	466	US-10-161-547-2	Sequence 2, Appli
10	34	61.8	605	US-10-156-761-14380	Sequence 14380, A
11	34	61.8	652	US-10-232-665-2	Sequence 2, Appli
12	34	61.8	652	US-10-232-665-4	Sequence 4, Appli
13	34	61.8	652	US-10-232-665-6	Sequence 6, Appli
14	34	61.8	653	US-10-232-665-8	Sequence 8, Appli
15	34	61.8	653	US-10-232-665-10	Sequence 10, Appli

16	34	61.8	653	US-10-232-665-12	Sequence 12, Appli
17	34	61.8	653	US-10-232-665-14	Sequence 14, Appli
18	34	61.8	653	US-10-232-665-16	Sequence 16, Appli
19	34	61.8	653	US-10-232-665-18	Sequence 18, Appli
20	34	61.8	653	US-10-232-665-20	Sequence 20, Appli
21	34	61.8	653	US-10-232-665-22	Sequence 22, Appli
22	34	61.8	653	US-10-232-665-24	Sequence 24, Appli
23	34	61.8	653	US-10-232-665-26	Sequence 26, Appli
24	34	61.8	653	US-10-232-665-28	Sequence 28, Appli
25	34	61.8	872	US-10-047-260-38	Sequence 38, Appli
26	34	61.8	1778	US-10-238-075-749	Sequence 749, Appli
27	33	60.0	170	US-09-746-783-16	Sequence 16, Appli
28	33	60.0	175	US-10-062-254-234	Sequence 234, Appli
29	33	60.0	467	US-09-875-573-17	Sequence 17, Appli
30	33	60.0	467	US-09-875-573-19	Sequence 19, Appli
31	33	60.0	467	US-09-875-573-21	Sequence 21, Appli
32	33	60.0	468	US-09-815-242-10201	Sequence 10201, A
33	33	60.0	468	US-09-815-242-13822	Sequence 13822, A
34	33	60.0	468	US-09-875-573-23	Sequence 23, Appli
35	33	60.0	468	US-09-875-573-25	Sequence 25, Appli
36	33	60.0	468	US-09-875-573-27	Sequence 27, Appli
37	33	60.0	468	US-09-875-573-29	Sequence 29, Appli
38	33	60.0	468	US-09-875-573-31	Sequence 31, Appli
39	33	60.0	468	US-09-875-573-33	Sequence 33, Appli
40	33	60.0	468	US-09-875-573-35	Sequence 35, Appli
41	33	60.0	468	US-09-875-573-37	Sequence 37, Appli
42	33	60.0	468	US-09-875-573-39	Sequence 39, Appli
43	33	60.0	468	US-09-875-573-41	Sequence 41, Appli
44	33	60.0	468	US-09-875-573-43	Sequence 43, Appli
45	33	60.0	473	US-09-815-242-5008	Sequence 5008, Ap

ALIGNMENTS

RESULT 1
US-10-205-942-4
; Sequence 4, Application US/10205942
; Publication No. US20030053990A1
; GENERAL INFORMATION:
; APPLICANT: University of No. US20030053990A1Ch Carolina-Chapel Hill
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samuelski, Richard J.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/10/205,942
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Adeno-associated virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(2271)
; OTHER INFORMATION: B19/AAV chimeric capsid coding sequence
US-10-205-942-4

Query Match 100.0%; Score 55; DB 15; Length 756;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKG2QYTDQ 10
|||
Db 597 NKG2QYTDQ 606

RESULT 2
US-10-188-586A-7
; Sequence 7, Application US/10188586A
; Publication No. US20030045438A1

GENERAL INFORMATION:
APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases From Gram Positive Organisms
FILE REFERENCE: GC389-US-D1
CURRENT APPLICATION NUMBER: US/10/188,586A
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 09/554,999
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: PCT/US98/27040
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: GB 9727471.6
PRIOR FILING DATE: 1997-12-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 233
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-188-586A-7

Query Match 63.6%; Score 35; DB 15; Length 233;
Best Local Similarity 60.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 206 NQKKKPYTDQ 215

RESULT 3
US-10-188-586A-2
Sequence 2, Application US/10188586A
Publication No. US20030045438A1
GENERAL INFORMATION:
APPLICANT: Estell, David A.
TITLE OF INVENTION: Proteases From Gram Positive Organisms
FILE REFERENCE: GC389-US-D1
CURRENT APPLICATION NUMBER: US/10/188,586A
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 09/554,999
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: PCT/US98/27040
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: GB 9727471.6
PRIOR FILING DATE: 1997-12-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-188-586A-2

Query Match 63.6%; Score 35; DB 15; Length 415;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 208 NQKKKPYTDQ 217

RESULT 4
US-10-205-979-49
Sequence 49, Application US/10205979
Publication No. US2003014761A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
APPLICANT: Abernethy, Nevlin
TITLE OF INVENTION: Compounds and Methods for the Modulation
TITLE OF INVENTION: of Immune Responses
FILE REFERENCE: 11000.1063U

CURRENT APPLICATION NUMBER: US/10/205,979
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/308,446
PRIOR FILING DATE: 2001-07-26
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(187)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-205-979-49

Query Match 61.8%; Score 34; DB 12; Length 187;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 161 DKSTHOYTDQ 170

RESULT 5
US-10-051-643-203
Sequence 203, Application US/10051643
Publication No. US20020197265A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Tan, Paul L. J.
TITLE OF INVENTION: Methods and Compounds for the Treatment
TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
FILE REFERENCE: 11000.1008C2
CURRENT APPLICATION NUMBER: US/10/051,643
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US09/156,181
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 08/996,624
PRIOR FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 203
LENGTH: 187
TYPE: PRT
ORGANISM: Mycobacterium vaccae
FEATURE:
NAME/KEY: UNSURE
LOCATION: (186)...(186)
US-10-051-643-203

Query Match 61.8%; Score 34; DB 14; Length 187;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 161 DKSTHOYTDQ 170

RESULT 6
US-10-161-547-14
Sequence 14, Application US/10161547
Publication No. US20030143216A1
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/10/161,547
CURRENT FILING DATE: 2002-06-03

PRIOR APPLICATION NUMBER: US/09/267,574
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-547-14

Query Match
Best Local Similarity 66.7%; Score 34; DB 12; Length 373;
Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
| | | | |
79 NFGTQKFTD 87

US-10-161-547-15
Sequence 15, Application US/10161547
Publication No. US20030143216A1
GENERAL INFORMATION:
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITTIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/10/161,547
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/267,574
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-547-15

Query Match
Best Local Similarity 66.7%; Score 34; DB 12; Length 373;
Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 NKGTQOYTD 9
| | | | |
79 NFGTQKFTD 87

RESULT 8
US-10-004-219B-10
Sequence 10, Application US/10004219B
Publication No. US20030087414A1
GENERAL INFORMATION:
APPLICANT: Macrozyme
APPLICANT: Aerts, Johannes M.F.G.
TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in
TITLE OF INVENTION: which mucus is involved or infection diseases
FILE REFERENCE: 2183-51360S
CURRENT APPLICATION NUMBER: US/10/004,219B
CURRENT FILING DATE: 2001-11-02
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: SITE
LOCATION: (1) ..(445)
OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

Query Match
Best Local Similarity 66.7%; Score 34; DB 15; Length 445;
Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
| | | | |
Db 79 NFGTQKFTD 87

RESULT 9
US-10-161-547-2
Sequence 2, Application US/10161547
Publication No. US20030143216A1
GENERAL INFORMATION:
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITTIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/10/161,547
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/267,574
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-547-2

Query Match
Best Local Similarity 66.7%; Score 34; DB 12; Length 466;
Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NKGTQOYTD 9
| | | | |
Db 100 NFGTQKFTD 108

RESULT 10
US-10-156-761-14380
Sequence 14380, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 14380
LENGTH: 605
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-14380

Query Match 61.8%; Score 34; DB 15; Length 605;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 10
| | | | |
Db 120 OGQOQYADQ 128

RESULT 11

US-10-232-665-2
; Sequence 2, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232,665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377,466
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-232-665-2

Query Match 61.8%; Score 34; DB 15; Length 652;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 235 KLTQOYTD 242

RESULT 12

US-10-232-665-4
; Sequence 4, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232,665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377,466
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-232-665-4

Query Match 61.8%; Score 34; DB 15; Length 652;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 235 KLTQOYTD 242

RESULT 13

US-10-232-665-6
; Sequence 6, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants

; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232,665
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377,466
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic or
; OTHER INFORMATION: non-naturally occurring amino acid sequence encoded by SEQ ID NO:
; FEATURE:
; NAME/KEY: PRT
; LOCATION: (1)..(652)
US-10-232-665-6

Query Match 61.8%; Score 34; DB 15; Length 652;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 235 KLTQOYTD 242

RESULT 14

US-10-232-665-8
; Sequence 8, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232,665
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377,466
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: non-naturally
; OTHER INFORMATION: occurring amino acid sequence encoded by SEQ ID NO: 7
; FEATURE:
; NAME/KEY: PRT
; LOCATION: (1)..(653)
; OTHER INFORMATION: amino acid sequence for Cry3Bb variant v11231 encoded by SEQ ID N
US-10-232-665-8

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
| | | | |
Db 236 KLTQOYTD 243

RESULT 15

US-10-232-665-10
; Sequence 10, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants

FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 10
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: non-naturally
OTHER INFORMATION: occurring amino acid sequence encoded by SEQ ID NO:9
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO:9
US-10-232-665-10
Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243
RESULT 16
US-10-232-665-12
Sequence 12, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 12
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: non-naturally
OTHER INFORMATION: occurring amino acid sequence encoded by SEQ ID NO:11
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(1653)
OTHER INFORMATION: amino acid sequence encoded by SEQ ID NO:11
US-10-232-665-12
Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243
RESULT 17
US-10-232-665-14
Sequence 14, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn

CURRENT APPLICATION NUMBER: US/10/232,665
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide encoded by SEQ ID NO:
OTHER INFORMATION:
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant v11231
US-10-232-665-14
Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243
RESULT 18
US-10-232-665-16
Sequence 16, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant v11231
US-10-232-665-16
Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243
RESULT 19
US-10-232-665-18
Sequence 18, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv1
US-10-232-665-18

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243

US-10-232-665-20
Sequence 20, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv1
US-10-232-665-20

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 KGTQOYTD 9
236 KLTQOYTD 243

RESULT 21
US-10-232-665-22
Sequence 22, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT

LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv2
US-10-232-665-22

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243

RESULT 22
US-10-232-665-24
Sequence 24, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: Cry3Bb1 variant 11231mv2
US-10-232-665-24

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTD 9
DB 236 KLTQOYTD 243

RESULT 23
US-10-232-665-37
Sequence 37, Application US/10232665
Publication No. US20030115630A1
GENERAL INFORMATION:
APPLICANT: Romano, Charles P.
TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
CURRENT APPLICATION NUMBER: US/10/232,665
PRIOR FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: US/09/377,466
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 653
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PRT
LOCATION: (1)..(653)
OTHER INFORMATION: variant Cry3Bb1 coding sequence encoding v11231
US-10-232-665-37

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 KGTQOYTD 9
| | | | |
Db 236 KLTQOYTD 243

RESULT 24

US-10-232-665-39
; Sequence 39, Application US/10232665
; Publication No. US20030115630A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
; CURRENT APPLICATION NUMBER: US/10/232,665
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US/09/377,466
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; ID NO 39
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PRT
; LOCATION: (1)..(653)
; OTHER INFORMATION: variant Cry3Bb1 coding sequence encoding v11231
US-10-232-665-39

Query Match 61.8%; Score 34; DB 15; Length 653;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 KGTQOYTD 9
| | | | |
Db 236 KLTQOYTD 243

RESULT 25

US-10-047-260-38
; Sequence 38, Application US/10047260
; Publication No. US20020164706A1
; GENERAL INFORMATION:
; APPLICANT: McCluskey, Michael
; APPLICANT: Larossa, Robert
; TITLE OF INVENTION: HIGH LEVEL PROMOTERS FROM CYANOBACTERIA
; FILE REFERENCE: C11715 US NA
; CURRENT APPLICATION NUMBER: US/10/047,260
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/264,925
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 872
; TYPE: PRT
; ORGANISM: Synechocystis sp. strain PCC6803
US-10-047-260-38

Query Match 61.8%; Score 34; DB 14; Length 872;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 2 KGTQOYTD 10
| | | | |
Db 146 RGTQOYTD 154

RESULT 26
US-10-238-075-749
; Sequence 749, Application US/10238075

; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat
; TITLE OF INVENTION: B.colli, and biological uses of these polynucleotides and of thei
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 749
; LENGTH: 1778
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-749

Query Match 61.8%; Score 34; DB 12; Length 1778;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 4 TQOYTDQ 10
| | | | |
Db 1682 TKOYTDQ 1688

RESULT 27

US-09-746-783-16
; Sequence 16, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; Lavallie, Edward R.
; Racie, Lisa A.
; Treacy, Maurice
; Spaulding, Vikki
; Agoastino, Michael J.
; Howes, Steven H.
; Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/746,783

FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milaeninc, Debra J.

REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 170 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-746-783-16

Query Match 60.0%; Score 33; DB 11; Length 170;
Best Local Similarity 50.0%; Pred. No. 87;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NKGTOQYTDQ 10
Db 109 NKSVENTYD 118

RESULT 28
US-10-062-254-234

Sequence 234, Application US/10062254
Publication No. US20020138882A1

GENERAL INFORMATION:

APPLICANT: Canoon, Edgar B

APPLICANT: Canoon, Rebecca E

APPLICANT: Falco, Saverio Carl

APPLICANT: Fang, Yiwen

APPLICANT: Hantke, Sabine S.

APPLICANT: Lee, Jian-Ming

APPLICANT: Li, Zhongsen

APPLICANT: Miao, Guo-Hua

APPLICANT: Morgante, Michele

APPLICANT: Niu, Xiping

APPLICANT: Odell, Joan

APPLICANT: Rafalski, Antoni

APPLICANT: Sakai, Hajime

APPLICANT: Zheng, Peizhong

APPLICANT: Zhu, Qun

TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/062,234

CURRENT FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 09/630,346

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/146511

PRIOR FILING DATE: 1999-07-30

PRIOR APPLICATION NUMBER: 60/156006

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: 60/156899

PRIOR FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: 60/157287

PRIOR FILING DATE: 1999-10-01

PRIOR APPLICATION NUMBER: 60/169767

PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 60/171054

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: 60/172958

PRIOR FILING DATE: 1999-12-21

PRIOR APPLICATION NUMBER: 60/171515

PRIOR FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: 60/173535

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 375

SOFTWARE: Microsoft Office 97

SEQ ID NO 234

LENGTH: 175

TYPE: PRT

ORGANISM: Trifolium aestivum

US-10-062-254-234

Query Match 60.0%; Score 33; DB 14; Length 175;

Best Local Similarity 75.0%; Pred. No. 89;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTOYTDQ 8

Db 77 NKGTRLYT 84

RESULT 29
US-09-875-573-17

Sequence 17, Application US/09875573

Patent No. US20020150902A1

GENERAL INFORMATION:

APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE

FILE REFERENCE: CHMED.001C1

CURRENT APPLICATION NUMBER: US/09/875,573

CURRENT FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/US99/29149

PRIOR FILING DATE: 1999-12-08

PRIOR APPLICATION NUMBER: US 60/111,493

PRIOR FILING DATE: 1998-12-08

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17

LENGTH: 467

TYPE: PRT

ORGANISM: Escherichia coli

US-09-875-573-17

Query Match 60.0%; Score 33; DB 10; Length 467;

Best Local Similarity 60.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOYTDQ 10

Db 258 NKGTKWTSQ 267

RESULT 30
US-09-875-573-19

Sequence 19, Application US/09875573

Patent No. US20020150902A1

GENERAL INFORMATION:

APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE

FILE REFERENCE: CHMED.001C1

CURRENT APPLICATION NUMBER: US/09/875,573

CURRENT FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/US99/29149

PRIOR FILING DATE: 1999-12-08

PRIOR APPLICATION NUMBER: US 60/111,493

PRIOR FILING DATE: 1998-12-08

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 19

LENGTH: 467

TYPE: PRT

ORGANISM: Escherichia coli

US-09-875-573-19

Query Match 60.0%; Score 33; DB 10; Length 467;

Best Local Similarity 60.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOYTDQ 10

Db 258 NKGTKWTSQ 267

RESULT 31
US-09-875-573-21

Sequence 21, Application US/09875573

Patent No. US20020150902A1

GENERAL INFORMATION:

APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE

FILE REFERENCE: CHMED.001C1

CURRENT APPLICATION NUMBER: US/09/875,573

CURRENT FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/US99/29149

PRIOR FILING DATE: 1999-12-08

PRIOR APPLICATION NUMBER: US 60/111,493

PRIOR FILING DATE: 1998-12-08

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 19

LENGTH: 467

TYPE: PRT

ORGANISM: Escherichia coli

US-09-875-573-21

Query Match 60.0%; Score 33; DB 10; Length 467;

Best Local Similarity 60.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOYTDQ 10

Db 258 NKGTKWTSQ 267

CURRENT APPLICATION NUMBER: US/09/875,573
CURRENT FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: PCT/US99/29149
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/111,493
PRIOR FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 467
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-21

Query Match 60.0%; Score 33; DB 10; Length 467;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
||||:|
258 NKGTGKWT SQ 267

RESULT 32
US-09-815-242-10201
Sequence 10201, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyckind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10201
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10201

Query Match 60.0%; Score 33; DB 9; Length 468;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
||||:|
Db 259 NKGTGKWT SQ 268

RESULT 33

US-09-815-242-13822
Sequence 13822, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyckind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13822
LENGTH: 468
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13822

Query Match 60.0%; Score 33; DB 9; Length 468;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
||||:|
Db 259 NKGTGKWT SQ 268

RESULT 34
US-09-875-573-23
Sequence 23, Application US/09875573
Patent No. US20020150902A1
GENERAL INFORMATION:
APPLICANT: Tarr, Phillip I.
TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
FILE REFERENCE: CMED.001C1
CURRENT APPLICATION NUMBER: US/09/875,573
CURRENT FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: PCT/US99/29149
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/111,493
PRIOR FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 468
TYPE: PRT
ORGANISM: Escherichia coli
US-09-875-573-23

Query Match 60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTDO 10

Db 259 NKGTKMTSQ 268

RESULT 35

US-09-875-573-25
; Sequence 25, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; SOFTWARE: FASTSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 25
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-875-573-25

US-09-875-573-25

Query Match

Best Local Similarity 60.0%; Score 33; DB 10; Length 468;
Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTDO 10

Db 259 NKGTKMTSQ 268

RESULT 36

US-09-875-573-27
; Sequence 27, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-875-573-27

US-09-875-573-27

Query Match

Best Local Similarity 60.0%; Score 33; DB 10; Length 468;
Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTDO 10

Db 259 NKGTKMTSQ 268

RESULT 37

US-09-875-573-29

Sequence 29, Application US/09875573

Patent No. US20020150902A1

GENERAL INFORMATION:

APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE

FILE REFERENCE: CHMED.001C1

CURRENT APPLICATION NUMBER: US/09/875,573

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/US99/29149

PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: US 60/111,493

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 29

LENGTH: 468

TYPE: PRT

ORGANISM: Escherichia coli

US-09-875-573-29

Query Match

Best Local Similarity 60.0%; Score 33; DB 10; Length 468;
Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTDO 10

Db 259 NKGTKMTSQ 268

RESULT 38

US-09-875-573-31

Sequence 31, Application US/09875573

Patent No. US20020150902A1

GENERAL INFORMATION:

APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE

FILE REFERENCE: CHMED.001C1

CURRENT APPLICATION NUMBER: US/09/875,573

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/US99/29149

PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: US 60/111,493

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 31

LENGTH: 468

TYPE: PRT

ORGANISM: Escherichia coli

US-09-875-573-31

Query Match

Best Local Similarity 60.0%; Score 33; DB 10; Length 468;
Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTOOYTDO 10

Db 259 NKGTKMTSQ 268

RESULT 39

US-09-875-573-33

Sequence 33, Application US/09875573

Patent No. US20020150902A1

GENERAL INFORMATION:

APPLICANT: Tarr, Phillip I.

TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE

FILE REFERENCE: CHMED.001C1

CURRENT APPLICATION NUMBER: US/09/875,573

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: PCT/US99/29149

PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: US 60/111,493

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 33

LENGTH: 468

TYPE: PRT

ORGANISM: Escherichia coli

US-09-875-573-33

Query Match


```

; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-875-573-33

```

```

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

Cy 1 NKGTQOYTDQ 10
    ||||:|
259 NKGTGKWTSTQ 268

```

```

RESULT 40
US-09-875-573-35
; Sequence 35, Application US/09875573
; Patent No. US20020150902A1
; GENERAL INFORMATION:
; APPLICANT: Tarr, Phillip I.
; TITLE OF INVENTION: POLYMORPHIC LOCI THAT DIFFERENTIATE
; FILE REFERENCE: CHMED.001C1
; CURRENT APPLICATION NUMBER: US/09/875,573
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US99/29149
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/111,493
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-875-573-35

```

```

Query Match          60.0%; Score 33; DB 10; Length 468;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

1 NKGTQOYTDQ 10
    ||||:|
259 NKGTGKWTSTQ 268

```

```

Search completed: August 20, 2003, 10:08:17
Job time : 40.6186 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 20, 2003, 09:19:07 ; Search time 11.2371 Seconds
(without alignments)
85.581 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKGTOQYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	781	1 VCPV19	coat protein VP1 -
2	38	69.1	3971	2 T44806	mycobactin synth
3	37	67.3	338	2 A53066	CCAAT enhancer-bin
4	35	63.6	169	2 D86801	prophage p13 prote
5	35	63.6	415	2 G69885	processing protein
6	34	61.8	106	2 A89872	conserved hypothet
7	34	61.8	375	2 T03325	gene 118 protein -
8	34	61.8	466	2 D84906	probable beta-keto
9	34	61.8	547	2 UC4519	heat-shock protein
10	34	61.8	548	2 C64076	chaperonin groEL -
11	34	61.8	652	2 I39811	parasporeal crystal
12	34	61.8	652	2 S10228	endopeptidase Clp
13	34	61.8	872	2 S76197	probable adhesin Z
14	34	61.8	1588	2 A86036	probable adhesin E
15	34	61.8	1588	2 H91188	hypothetical prote
16	33	60.0	163	2 T48609	hypothetical prote
17	33	60.0	223	2 T19194	hypothetical prote
18	33	60.0	224	2 T19199	hypothetical prote
19	33	60.0	238	2 S31779	trypsin (BC 3.4.21
20	33	60.0	246	2 H82553	2'-demethylubiquin
21	33	60.0	310	2 T02202	2-hydroxyisoflavo
22	33	60.0	375	2 H90786	hypothetical prote
23	33	60.0	375	2 H85646	hypothetical prote
24	33	60.0	443	2 I40708	phosphogluconate d
25	33	60.0	445	2 I40629	phosphogluconate d
26	33	60.0	445	2 I40681	phosphogluconate d
27	33	60.0	445	2 I40682	phosphogluconate d
28	33	60.0	445	2 I40684	phosphogluconate d
29	33	60.0	445	2 I40685	phosphogluconate d

30	33	60.0	445	2 I40709	phosphogluconate d
31	33	60.0	468	1 DECCGC	phosphogluconate d
32	33	60.0	468	1 S04397	phosphogluconate d
33	33	60.0	468	2 I62463	phosphogluconate d
34	33	60.0	468	2 I62465	phosphogluconate d
35	33	60.0	468	2 D56146	phosphogluconate d
36	33	60.0	468	2 I84555	phosphogluconate d
37	33	60.0	468	2 I41249	phosphogluconate d
38	33	60.0	468	2 I41250	phosphogluconate d
39	33	60.0	468	2 P90982	gluconate-6-phosph
40	33	60.0	468	2 D85828	gluconate-6-phosph
41	33	60.0	468	2 AE0765	phosphogluconate d
42	33	60.0	469	2 A10187	phosphogluconate d
43	33	60.0	474	2 E95043	hypothetical prote
44	33	60.0	474	2 G97913	phosphogluconate d
45	33	60.0	498	2 B90456	hypothetical prote

ALIGNMENTS

RESULT 1

VCPV19
coat protein VP1 - parvovirus B19 (strain Au)

C:Species: parvovirus B19

A:Note: host Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C:Accession: A24239

R:Shade, R.O.; Blundell, M.C.; Cotmore, S.F.; Tattersall, P.; Atwell, C.R.

J. Virol. 58, 921-936, 1986

A:Title: Nucleotide sequence and genome organization of human parvovirus B19 isolated fr

A:Reference number: A24239; MUID:86200451; PMID:3701931

A:Accession: A24239

A:Molecule type: DNA

A:Residues: 1-781 <SHA>

A:Cross-references: EMBL:M13178; NID:G333375; PIDN:AA66867.1; PID:G333377

A:Superfamily: parvovirus coat protein

C:Keywords: coat protein

Query Match 100.0%; Score 55; DB 1; Length 781;

Best Local Similarity 100.0%; Pred. No. 0.0061;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTOQYTDQ 10
DB 622 NKGTOQYTDQ 631

RESULT 2

T44806
mycobactin synthetase chain myca (imported) - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Nov-2000

C:Accession: T44806

R:Dultman, E.H.; Hamoen, L.W.; Rembold, M.; Venema, G.; Seitz, H.; Saenger, W.; Bernhard,

Proc. Natl. Acad. Sci. U.S.A. 96, 13294-13299, 1999

A:Title: The mycobactin synthetase of Bacillus subtilis ATCC6633: A multifunctional hy

A:Reference number: Z22848; MUID:20027541; PMID:10557314

A:Accession: T44806

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3971 <DUI>

A:Cross-references: EMBL:AF184956; NID:G6449053; PIDN:AAF08795.1; PID:G6449055

A:Experimental source: strain ATCC6633

A:Genetics: myca

C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase ;

C:Keywords: antibiotic biosynthesis; carrier protein

F:56-548/Domain: acetate-CoA ligase homology <ACLI>

F:581-649/Domain: acyl carrier protein homology <ACPI>

F:690-1090/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>

F:1290-1361/Domain: acyl carrier protein homology <ACFP>

F:2408-2476/Domain: acyl carrier protein homology <ACFP>

F:2986-3427/Domain: acetate-CoA ligase homology <AC12>
F:3445-3513/Domain: acyl carrier protein homology <ACP4>

Query Match 69.1%; Score 38; DB 2; Length 3971;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGTQOQYTDQ 10
|:|:|:|:|:
DB 1496 KGSQOYTDQ 1504

RESULT 3

CCAT enhancer-binding protein - California sea hare

A:Accession: A53066
N:Alternate names: C/EBP
C:Species: Aplysia californica (California sea hare)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A53066
R:Alberini, C.M.; Ghirardi, M.; Metz, R.; Kandel, E.R.

File: C/EBP is an immediate-early gene required for the consolidation of long-term fa
Reference number: A53066; MUID:94185169; PMID:8137425

A:Accession: A53066
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-338 <ALB>
A:Cross-references: GB:U00994; NID:G392976; PIDN:AAA18286.1; PID:9487963

Query Match 67.3%; Score 37; DB 2; Length 338;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOQYTDQ 10
|:|:|:|:|:
DB 258 KGTQETVDK 266

RESULT 4

prophage p13 protein 39 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D86801
R:Boletín, A.; Winkler, P.; Mauger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 733-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: D86801
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <STO>
A:Cross-references: GB:AE005176; PID:G12724400; PIDN:AAK05510.1; GSPDB:GN00146
A:Experimental source: strain IL1403
A:Gene: p1339

Query Match 63.6%; Score 35; DB 2; Length 169;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQOQYTDQ 10
|:|:|:|:|:
DB 60 NNSQOQYADQ 69

RESULT 5

processsing protease homolog ymfH - Bacillus subtilis

C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: G69885
R:Kunst, F.; Ogatawara, N.; Moszer, I.; Alberini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Galler
leth, J.; Harwood, C.R.; Henauc, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koster, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:96044033; PMID:9364377

A:Accession: G69885
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-415 <KUN>

A:Cross-references: GB:Z99112; GB:AL009126; NID:G2633902; PIDN:CAB13559.1; PID:G2634058
A:Experimental source: strain 168

C:Gene: ymfH
C:Superfamily: mitochondrial processing peptidase alpha chain

Query Match 63.6%; Score 35; DB 2; Length 415;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQOQYTDQ 10
|:|:|:|:|:
DB 208 NQKKPYTDQ 217

RESULT 6

conserved hypothetical protein SA0887 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: A89872
R:Kuroda, M.; Ohta, T.; Uchiyama, T.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cuí, L.; Oguci
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogatawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89872
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <KUN>
A:Cross-references: GB:BA000016; PID:G13700836; PIDN:BAK42132.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Gene: SA0887

Query Match 61.8%; Score 34; DB 2; Length 106;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOQYTD 9
|:|:|:|:|:
DB 45 KGTQOQYVN 52

RESULT 7

gene 118 protein - Lactococcus phage b11170

C:Species: Lactococcus phage b11170
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 04-Mar-2000
C:Accession: T03325
R:Critt-De Coq, A.M.; Casselin, B.; Commissaire, J.; Anba, J.; Kyriakidis, S.; Chopin, M
submitted to the EMBL Data Library, June 1997
A:Description: Sequence and organization of the lactococcal isometric b11170 phage genom

A:Reference number: Z14903
A:Accession: T03325
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-375 <CRU>
A:Cross-references: EMBL:AF009650; NID:g3282260; PIDD:AA27197.1; PID:g3282278
C:Genetics:
A:Gene: 118
C:Superfamily: Lactococcus phage b1170 gene 118 protein

Query Match 61.8%; Score 34; DB 2; Length 375;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 NKGTQOYTD 9
|||
Db 250 NKATDDYTD 258

UT 8
06
Probable beta-ketoacyl-CoA synthase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84906
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.
euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84906
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <STO>
A:Cross-references: GB:AE002093; NID:g3831446; PIDD:AA269929.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2946720
A:Map position: 2

Query Match 61.8%; Score 34; DB 2; Length 466;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 KGTQOYTDQ 10
|||
86 KGAQYTD 94

RESULT 9
JC4519
heat-shock protein groEL - Pasteurella multocida
C:Species: Pasteurella multocida
C>Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 26-Aug-1999
C:Accession: JC4519
R:Love, B.C.; Hansen, L.M.; Hirsch, D.C.
Gene 166, 179-180, 1995
A:Title: Cloning and sequence of the groEL heat-shock operon of Pasteurella multocida.
A:Reference number: JC4518; MUID:96105224; PMID:8529887
A:Accession: JC4519
A:Molecule type: DNA
A:Residues: 1-547 <LOV>
A:Cross-references: GB:U0165; NID:g1144300; PIDD:AAA84916.1; PID:g1144302
A:Experimental source: serotype A:3
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; stress-induced protein

Query Match 61.8%; Score 34; DB 2; Length 547;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NKGTQOYTD 9
|||
Db 479 NAGTEOYGD 487

RESULT 10
C64076
chaperonin groEL - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999
C:Accession: C64076
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
J.; Goodyear, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodet, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542600
A:Accession: C64076
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-548 <TIGR>
A:Cross-references: GB:U32736; GB:I42023; NID:g1573519; PIDD:AA22201.1; PID:g1573528; T
C:Genetics:
A:Gene: groEL
C:Superfamily: chaperonin groEL
C:Keywords: heat shock; molecular chaperone; stress-induced protein

Query Match 61.8%; Score 34; DB 2; Length 548;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NKGTQOYTD 9
|||
Db 480 NAGTEOYGD 488

RESULT 11
I39811
paraspore crystal protein cry3Bb1 - Bacillus thuringiensis
N:Alternate names: paraspore crystal protein cry11B2
C:Species: Bacillus thuringiensis
C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 01-Dec-2000
C:Accession: I39811
R:Donovan, W.P.; Rupa, M.J.; Slaney, A.C.; Malvar, T.; Gawron-Burke, M.C.; Johnson, T.B
Appl. Environ. Microbiol. 58, 3921-3927, 1992
A:Title: Characterization of two genes encoding Bacillus thuringiensis insecticidal crys
A:Reference number: I39811; MUID:93119147; PMID:1476436
A:Accession: I39811
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-652 <RES>
A:Cross-references: GB:M89794; NID:g142729; PIDD:AAA22334.1; PID:g142730
C:Genetics:
A:Gene: cry11B2
C:Superfamily: paraspore crystal protein

Query Match 61.8%; Score 34; DB 2; Length 652;
Best Local Similarity 87.5%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 KGTQOYTD 9
|||
Db 235 KLTQOYTD 242

RESULT 12
S10228
paraspore crystal protein cry3Ba1 - Bacillus thuringiensis (fragment)
N:Alternate names: coleopterac-active paraspore crystal protein; delta-endotoxin
C:Species: Bacillus thuringiensis
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Dec-2000
C:Accession: S10228

R:Sick, A.; Gaertner, F.; Wong, A.
Nucleic Acids Res. 18, 1305, 1990
A>Title: Nucleotide sequence of a coleopteran-active toxin gene from a new isolate of *B.*
A:Reference number: S10228; PMID:90206811; PMID:2320431
A:Accession: S10228
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-659 <SIC>
A:Cross-references: EMBL:X17123; NID:g40258; PIDN:CA34983.1; PID:g40259
A:Gene: cryIIIB
A:Superfamily: parasporal crystal protein
A:Keywords: delta-endotoxin; toxin

Query Match 61.8%; Score 34; DB 2; Length 659;
Best Local Similarity 87.5%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTDQ 9
|:|||||
243 KLTQOYTD 250

RESULT 13

S76197
endopeptidase Clp ATP-binding chain B1 - *Synechocystis* sp. (strain PCC 6803)
N:Alternate names: ATP-dependent Clp protease regulatory chain; protein b1r1641
M:Comments: adenosinetriphosphatase (EC 3.6.1.3)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 19-Jan-2001
C:Accession: S76197

R:Kanaoka, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
B.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76197
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-872 <KAN>
A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BA18456.1; PID:g165354
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: clpB1

C:Function:
A:Description: allows clpB to hydrolyze polypeptides and proteins, probably by a chaperon
activity; ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller
peptides; endopeptidase Clp ATP-binding chain
A:Keywords: ATP; hydrolase; molecular chaperone; nucleotide binding; P-loop
F:208-215/Region: nucleotide-binding motif A (P-loop)
F:276-280/Region: nucleotide-binding motif B
F:611-618/Region: nucleotide-binding motif A (P-loop)
F:679-683/Region: nucleotide-binding motif B
F:214/Binding site: ATP (lys) #status predicted
F:617/Binding site: ATP (lys) #status predicted

Query Match 61.8%; Score 34; DB 2; Length 872;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KGTQOYTDQ 10
|:|||||
Db 146 KGTQOYTDQ 154

RESULT 14

A86036
probable adhesin 25029 (imported) - *Escherichia coli* (strain O157:H7, substrain EDL933)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: A86036

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouets, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A86036
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1588 <STO>
A:Cross-references: GB:AE005174; NID:g12518349; PIDN:AG58749.1; GSPDB:GN00145; UMGF:Z50
A:Experimental source: strain O157:H7, substrain EDL933
A:Genetics:
A:Gene: 25029

Query Match 61.8%; Score 34; DB 2; Length 1588;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TQOYTDQ 10
|:|||||
Db 1492 TKOYTDQ 1498

RESULT 15

H9118
probable adhesin EC84480 [similarity] - *Escherichia coli* (strain O157:H7, substrain RMD
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C:Accession: H9118
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawara, N.; Yasunaga, T.; Kuahara, S.; Shiba, T.; Hatlori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno-

A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H9118
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1588 <HAV>
A:Cross-references: GB:BA000007; PIDN:BA37903.1; PID:g13363955; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
A:Genetics:
A:Gene: EC84480

Query Match 61.8%; Score 34; DB 2; Length 1588;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TQOYTDQ 10
|:|||||
Db 1492 TKOYTDQ 1498

RESULT 16

T48609
hypothetical protein F18022.150 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48609
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Len
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24493
A:Accession: T48609
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163 <BEV>
A:Cross-references: EMBL:DB
A:Experimental source: cultivar Columbia; BAC clone F18022
C:Genetics:
A:Map position: 5
A:introns: 92/1
A:Note: F18022.150

Query Match 60.0%; Score 33; DB 2; Length 163;

Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 NKGTQOYTD 10
|||||
Db 21 NKGTQOQDQ 30

RESULT 17

T19194

hypochemical protein C11B4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T19194

R:Mortimore, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19087

A:Accession: T19194

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Cross-references: EMBL:Z81015; PIDN:CA802659.1; GSPDB:GN00028; CESP:C11B4.1

A:Experimental source: clone C11B4

C:Genetics:

A:Gene: CESP:C11B4.1

A:Map position: X

A:Introns: 20/1; 44/3; 73/1; 154/3

C:Superfamily: glutathione peroxidase

Query Match

60.0%; Score 33; DB 2; Length 223;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TOOYTD 9
|||||
Db 75 TOOYTD 80

RESULT 18

T19190

hypochemical protein C11B4.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T19190

R:Mortimore, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19087

A:Accession: T19190

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-224 <MIL>

A:Cross-references: EMBL:Z81015; PIDN:CA802655.1; GSPDB:GN00028; CESP:C11B4.2

A:Experimental source: clone C11B4

C:Genetics:

A:Gene: CESP:C11B4.2

A:Map position: X

A:Introns: 22/1; 74/1; 155/3

C:Superfamily: glutathione peroxidase

Query Match

60.0%; Score 33; DB 2; Length 224;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TOOYTD 9
|||||
Db 76 TOOYTD 81

RESULT 19

S31779

trypsin (EC 3.4.21.4) III precursor - Atlantic salmon (fragment)

C:Species: Salmo salar (Atlantic salmon)

C:Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999

C:Accession: S66657; S31779
R:Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A:Title: Molecular cloning and characterization of anionic and cationic variants of tryp
A:Reference number: S66657; MUID:9603908; PMID:7556223
A:Accession: S66657

A:Molecule type: mRNA

A:Residues: 1-238 <MAL>

A:Cross-references: EMBL:X70074; NID:964387; PIDN:CA449679.1; PID:964388

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-7/Domains: signal sequence (fragment) #status predicted <SIG>

F:8-15/Domains: activation peptide #status predicted <APT>

F:16-238/Product: trypsin III #status predicted <MAT>

F:16-231/Domains: trypsin homology <TRY>

F:22-152,40-56,124-225,131-198,163-177,188-212/Disulfide bonds: #status predicted

F:35,99,192/Active site: His, Asp, Ser #status predicted

Query Match

60.0%; Score 33; DB 2; Length 238;

Best Local Similarity 55.6%; Pred. No. 49;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NKGTQOYTD 9
|||||
Db 73 NKGTQOYTD 81

RESULT 20

H82553

3-demethylubiquinone-9-3-methyltransferase XF2471 [imported] - Xylella fastidiosa (strain

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: H82553

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: H82553

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-246 <SIM>

A:Cross-references: GB:AE004055; GB:AE003849; NID:99107661; PIDN:JAF85269.1; GSPDB:GN0012

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Agencio, M.; Alvarenga, R.; Al

Bioness, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H

as-Neto, E.; Docena, C.; El-Dorfi, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohne

J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.B.; Laigre

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki

M.; Tsuchioka, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zé

A:Reference number: A59328

A:Accession: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2471

C:Superfamily: 3-demethylubiquinone-9-3-O-methyltransferase; bioc homology

Query Match 60.0%; Score 33; DB 2; Length 246;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KGTQOYTD 9
|||||
Db 182 KGTQOYTD 189

RESULT 21

T02202

RESULT 26
140681
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Citrobacter diversus (F
N/Alternate names: 6-phosphoglucuronate dehydrogenase
C/Species: Citrobacter diversus
C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 11-Jun-1999
C/Accession: 140681; 140683; 140686
R/Nelson, K.; Selander, R.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 10227-10231, 1994
A/Title: Intergenic transfer and recombination of the 6-phosphoglucuronate dehydrogenase
A/Reference number: 140629; PMID:95024018; PMID:7937867
A/Accession: 140681
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-445 <RES>
A/Cross-references: EMBL:U14424; NID:9538161; PIDN:AAC43774.1; PID:9538162
A/Experimental source: strain CT19
A/Accession: 140683
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-445 <RES>
A/Cross-references: EMBL:U14427; NID:9538167; PIDN:AAC43776.1; PID:9538168
A/Experimental source: strain CT4
A/Accession: 140686
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-445 <RES>
A/Cross-references: EMBL:U14432; NID:9538177; PIDN:AAC43779.1; PID:9538178
A/Experimental source: strain CT9
C/Genetics:
A/Gene: gnd
C/Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de
C/Keywords: oxidoreductase
F:1-274/Domain: 3-hydroxyisobutyrate dehydrogenase homology (fragment) <HIB>

Query Match 60.0%; Score 33; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Cy 1 NKGTQQTDD 10
|||::|
Db 248 NKGTGKWTSS 257
|||::|
RESULT 27
140682
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Citrobacter diversus (F
N/Alternate names: 6-phosphoglucuronate dehydrogenase
C/Species: Citrobacter diversus
C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 11-Jun-1999
C/Accession: 140682
R/Nelson, K.; Selander, R.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 10227-10231, 1994
A/Title: Intergenic transfer and recombination of the 6-phosphoglucuronate dehydrogenase
A/Reference number: 140629; PMID:95024018; PMID:7937867
A/Accession: 140682
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-445 <RES>
A/Cross-references: EMBL:U14425; NID:9538163; PIDN:AAC43775.1; PID:9538164
C/Genetics:
A/Gene: gnd
C/Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de
C/Keywords: oxidoreductase
F:1-274/Domain: 3-hydroxyisobutyrate dehydrogenase homology (fragment) <HIB>

Query Match 60.0%; Score 33; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Cy 1 NKGTQQTDD 10
|||::|
Db 248 NKGTGKWTSS 257
|||::|

RESULT 28
140684
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Citrobacter diversus (F
N/Alternate names: 6-phosphoglucuronate dehydrogenase
C/Species: Citrobacter diversus
C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 11-Jun-1999
C/Accession: 140684
R/Nelson, K.; Selander, R.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 10227-10231, 1994
A/Title: Intergenic transfer and recombination of the 6-phosphoglucuronate dehydrogenase
A/Reference number: 140629; PMID:95024018; PMID:7937867
A/Accession: 140684
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-445 <RES>
A/Cross-references: EMBL:U14428; NID:9538169; PIDN:AAC43777.1; PID:9538170
C/Genetics:
A/Gene: gnd
C/Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de
C/Keywords: oxidoreductase
F:1-274/Domain: 3-hydroxyisobutyrate dehydrogenase homology (fragment) <HIB>

Query Match 60.0%; Score 33; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Cy 1 NKGTQQTDD 10
|||::|
Db 248 NKGTGKWTSS 257
|||::|

RESULT 29
140685
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Citrobacter diversus (F
N/Alternate names: 6-phosphoglucuronate dehydrogenase
C/Species: Citrobacter diversus
C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 11-Jun-1999
C/Accession: 140685
R/Nelson, K.; Selander, R.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 10227-10231, 1994
A/Title: Intergenic transfer and recombination of the 6-phosphoglucuronate dehydrogenase
A/Reference number: 140629; PMID:95024018; PMID:7937867
A/Accession: 140685
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-445 <RES>
A/Cross-references: EMBL:U14429; NID:9538171; PIDN:AAC43778.1; PID:9538172
C/Genetics:
A/Gene: gnd
C/Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de
C/Keywords: oxidoreductase
F:1-274/Domain: 3-hydroxyisobutyrate dehydrogenase homology (fragment) <HIB>

Query Match 60.0%; Score 33; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Cy 1 NKGTQQTDD 10
|||::|
Db 248 NKGTGKWTSS 257
|||::|

RESULT 30
140709
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Citrobacter freundii (F
C/Species: Citrobacter freundii
C/Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 11-Jun-1999
C/Accession: 140709
R/Nelson, K.; Selander, R.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 10227-10231, 1994
A/Title: Intergenic transfer and recombination of the 6-phosphoglucuronate dehydrogenase

A:Reference number: 140629; MUID:95024018; PMID:7937867
A:Accession: 140709
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-445 <RES>
A:Cross-references: EMBL:U14466; NID:G540104; PIDN:AA03814.1; PID:G540105
C:Genetics:
A:Gene: gnd
C:Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate dehydrogenase (decarboxylating)
C:Keywords: oxidoreductase
F.1-274/Domain: 3-hydroxyisobutyrate dehydrogenase homology (fragment) <HID>

Query Match 60.0%; Score 33; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
DB 248 NKGTKWTSQ 257
|||||
|||||

BLAST 31

phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Escherichia coli (strain ATCC 8739)

N:Alternate names: gluconate-6-phosphate dehydrogenase
C:Species: Escherichia coli
C:Date: 28-Feb-1988 #sequence_revision 23-Jan-1998 #text_change 01-Mar-2002
C:Accession: D64966; 162464; 162465; A00360; I41248; I41246; I41243; I41244; I41245; I41247; I41249; I41250; I41251; I41252; I41253; I41254; I41255; I41256; I41257; I41258; I41259; I41260; I41261; I41262; I41263; I41264; I41265; I41266; I41267; I41268; I41269; I41270; I41271; I41272; I41273; I41274; I41275; I41276; I41277; I41278; I41279; I41280; I41281; I41282; I41283; I41284; I41285; I41286; I41287; I41288; I41289; I41290; I41291; I41292; I41293; I41294; I41295; I41296; I41297; I41298; I41299; I41300; I41301; I41302; I41303; I41304; I41305; I41306; I41307; I41308; I41309; I41310; I41311; I41312; I41313; I41314; I41315; I41316; I41317; I41318; I41319; I41320; I41321; I41322; I41323; I41324; I41325; I41326; I41327; I41328; I41329; I41330; I41331; I41332; I41333; I41334; I41335; I41336; I41337; I41338; I41339; I41340; I41341; I41342; I41343; I41344; I41345; I41346; I41347; I41348; I41349; I41350; I41351; I41352; I41353; I41354; I41355; I41356; I41357; I41358; I41359; I41360; I41361; I41362; I41363; I41364; I41365; I41366; I41367; I41368; I41369; I41370; I41371; I41372; I41373; I41374; I41375; I41376; I41377; I41378; I41379; I41380; I41381; I41382; I41383; I41384; I41385; I41386; I41387; I41388; I41389; I41390; I41391; I41392; I41393; I41394; I41395; I41396; I41397; I41398; I41399; I41400; I41401; I41402; I41403; I41404; I41405; I41406; I41407; I41408; I41409; I41410; I41411; I41412; I41413; I41414; I41415; I41416; I41417; I41418; I41419; I41420; I41421; I41422; I41423; I41424; I41425; I41426; I41427; I41428; I41429; I41430; I41431; I41432; I41433; I41434; I41435; I41436; I41437; I41438; I41439; I41440; I41441; I41442; I41443; I41444; I41445; I41446; I41447; I41448; I41449; I41450; I41451; I41452; I41453; I41454; I41455; I41456; I41457; I41458; I41459; I41460; I41461; I41462; I41463; I41464; I41465; I41466; I41467; I41468; I41469; I41470; I41471; I41472; I41473; I41474; I41475; I41476; I41477; I41478; I41479; I41480; I41481; I41482; I41483; I41484; I41485; I41486; I41487; I41488; I41489; I41490; I41491; I41492; I41493; I41494; I41495; I41496; I41497; I41498; I41499; I41500; I41501; I41502; I41503; I41504; I41505; I41506; I41507; I41508; I41509; I41510; I41511; I41512; I41513; I41514; I41515; I41516; I41517; I41518; I41519; I41520; I41521; I41522; I41523; I41524; I41525; I41526; I41527; I41528; I41529; I41530; I41531; I41532; I41533; I41534; I41535; I41536; I41537; I41538; I41539; I41540; I41541; I41542; I41543; I41544; I41545; I41546; I41547; I41548; I41549; I41550; I41551; I41552; I41553; I41554; I41555; I41556; I41557; I41558; I41559; I41560; I41561; I41562; I41563; I41564; I41565; I41566; I41567; I41568; I41569; I41570; I41571; I41572; I41573; I41574; I41575; I41576; I41577; I41578; I41579; I41580; I41581; I41582; I41583; I41584; I41585; I41586; I41587; I41588; I41589; I41590; I41591; I41592; I41593; I41594; I41595; I41596; I41597; I41598; I41599; I41600; I41601; I41602; I41603; I41604; I41605; I41606; I41607; I41608; I41609; I41610; I41611; I41612; I41613; I41614; I41615; I41616; I41617; I41618; I41619; I41620; I41621; I41622; I41623; I41624; I41625; I41626; I41627; I41628; I41629; I41630; I41631; I41632; I41633; I41634; I41635; I41636; I41637; I41638; I41639; I41640; I41641; I41642; I41643; I41644; I41645; I41646; I41647; I41648; I41649; I41650; I41651; I41652; I41653; I41654; I41655; I41656; I41657; I41658; I41659; I41660; I41661; I41662; I41663; I41664; I41665; I41666; I41667; I41668; I41669; I41670; I41671; I41672; I41673; I41674; I41675; I41676; I41677; I41678; I41679; I41680; I41681; I41682; I41683; I41684; I41685; I41686; I41687; I41688; I41689; I41690; I41691; I41692; I41693; I41694; I41695; I41696; I41697; I41698; I41699; I41700; I41701; I41702; I41703; I41704; I41705; I41706; I41707; I41708; I41709; I41710; I41711; I41712; I41713; I41714; I41715; I41716; I41717; I41718; I41719; I41720; I41721; I41722; I41723; I41724; I41725; I41726; I41727; I41728; I41729; I41730; I41731; I41732; I41733; I41734; I41735; I41736; I41737; I41738; I41739; I41740; I41741; I41742; I41743; I41744; I41745; I41746; I41747; I41748; I41749; I41750; I41751;

A>Title: Comparative nucleotide sequence analysis of growth-rate-regulated *gnd* alleles
 A:Reference number: 141243; MUID:88086897; PMID:3275621
 A:Accession: 141246
 A>Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-125 <RB4>
 A:Cross-references: GB:M18960; NID:g146239; PIDN:AAA33922.1; PID:g146240
 A:Experimental source: strain 558
 A:Accession: 141243
 A>Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-125 <RE2>
 A:Cross-references: GB:M18956; NID:g146233; PIDN:AAA33919.1; PID:g146234
 A:Experimental source: strain B/E
 A:Accession: 141244
 A>Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-125 <RE3>
 A:Cross-references: GB:M18957; NID:g146235; PIDN:AAA33920.1; PID:g146236
 A:Experimental source: strain 740
 A:Accession: 141245
 A>Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-130, 'V', 32-36, 'D', 38-53, 'F', 55-125 <RE6>
 A:Cross-references: GB:M18958; NID:g146237; PIDN:AAA33921.1; PID:g146238
 A:Experimental source: strain 567
 A:Accession: 141247
 A>Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-48, 'Q', 50-125 <RE7>
 A:Cross-references: GB:M18961; NID:g146241; PIDN:AAA33923.1; PID:g146242
 A:Experimental source: strain 745
 R:Marolda, C.L.; Valvano, M.A.
 J. Bacteriol. 175, 148-158, 1993
 A>Title: Identification, expression, and DNA sequence of the GDP-mannose biosynthesis *gnd*
 A:Reference number: AA0630; MUID:93106949; PMID:7677991
 A:Accession: EA0630
 A:Molecule type: DNA
 A:Residues: 1-38, 'Q', 40-122, 'YRY', <MAR>
 A:Cross-references: GB:L04596; NID:95524160; PIDN:AAC27540.1; PID:9415625
 A:Experimental source: strain WM187 (Escherichia coli O7:K1)
 A>Note: sequence extracted from NCBI backbone (NCBI:121128, NCBI:P.121133)
 A>Note: the final three residues of this fragment differ by an apparent frameshift error
 R:Bieleric, M.; Feutrier, J.Y.; Reeves, P.R.
 J. Bacteriol. 173, 3894-3900, 1991
 A>Title: Nucleotide sequences of the *gnd* genes from nine natural isolates of *Escherichia coli*.
 A:Reference number: 154956; MUID:91267957; PMID:2050640
 A:Accession: 154956
 A>Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-51, 'D', 53-54, 'F', 56-169, 'F', 171-210, 'S', 212-339, 'S', 331-349, 'A', 351-368, 'R'
 A:Cross-references: GB:M63821; NID:g147486; PIDN:AAA44488.1; PID:g147487
 A:Experimental source: strain EOCOR10
 A:Comment: This phosphogluconate dehydrogenase catalyzes the oxidative decarboxylation of
 ce shunt.
 C:Genetics:
 A:Gene: *gnd*
 A:Map position: 44 min
 C:Superfamily: phosphoglucate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de
 C:Keywords: NADP; oxidative decarboxylation; oxidoreductase; pentose phosphate pathway
 P:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <H1B>
 Query Match 60.0%; Score 33; DB 1; Length 468;
 Best Local Similarity 60.0%; Pred. No. 1e+02;
 Matches 6; conservative 2; Mismatches 2; Indels 0; Gaps 0;
 1 NKGTCQYTDQ 10
 Db 259 NKGTCQYTDQ 268

RESULT 32

```
Qy      1  NKGTDQYTDQ  10
          |||| :|| |
Db      259 NKGTDKWTSD  268
```

RESULT 32

S04397
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S04397; S15315
R:Reeves, P.; Stevenson, G.
Mol. Gen. Genet. 217, 182-184, 1989
A:Title: Cloning and nucleotide sequence of the Salmonella typhimurium LT2 gnd gene and
A:Reference number: S04397; MUID:89364685; PMID:2671649
A:Accession: S04397
A:Molecule type: DNA
A:Residues: 1-468 <RES>
R:Jiang, X.M.; Neal, B.; Santiago, F.; Lee, S.J.; Romana, L.K.; Reeves, P.R.
Mol. Microbiol. 5, 695-713, 1991
A:Title: Structure and sequence of the rfb (O antigen) gene cluster of Salmonella serovar
A:Reference number: S15296; MUID:91260454; PMID:1710759
A:Accession: S15315
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-57 <JIA>
A:Gene: gnd
C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de
C:Keywords: oxidoreductase; pentose phosphate pathway
F:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 60.0%; Score 33; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
|||::||
Db 259 NKGTKWTSQ 268
RESULT 33
162463
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Escherichia coli (strai
C:Species: Escherichia coli
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 11-Jun-1999
C:Accession: 162463
R:dykhuizen, D.E.; Green, L.
J. Bacteriol. 173, 7257-7268, 1991
A:Title: Recombination in Escherichia coli and definition of biological species.
A:Reference number: 141249; MUID:92041624; PMID:1938920
A:Accession: 162463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-468 <RES>
C:Cross-references: GB:M64329; NID:G146937; PIDN:AAA24207.1; PID:G146938
C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate-de
C:Keywords: oxidoreductase
F:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 60.0%; Score 33; DB 2; Length 468;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
|||::||
Db 259 NKGTKWTSQ 268
RESULT 34
162465
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Escherichia coli (strai
C:Species: Escherichia coli
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 11-Jun-1999
C:Accession: 162465
R:dykhuizen, D.E.; Green, L.
J. Bacteriol. 173, 7257-7268, 1991
A:Title: Recombination in Escherichia coli and definition of biological species.
A:Reference number: 141249; MUID:92041624; PMID:1938920

A:Accession: 162465
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-468 <RES>
C:Cross-references: GB:M64331; NID:G146941; PIDN:AAA24209.1; PID:G146942
C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate dei
C:Keywords: oxidoreductase
F:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 60.0%; Score 33; DB 2; Length 468;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
|||::||
Db 259 NKGTKWTSQ 268
RESULT 35
D56146
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
C:Accession: D56146
R:Arakawa, Y.; Macharotayankun, R.; Nagatsuka, T.; Ito, H.; Kato, N.; Ohta, M.
J. Bacteriol. 177, 1788-1796, 1995
A:Title: Genomic organization of the Klebsiella pneumoniae cps region responsible for sei
A:Reference number: A56146; MUID:95204345; PMID:7836702
A:Accession: D56146
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <ARA>
C:Cross-references: GB:D21242; NID:G747654; PIDN:BA04786.1; PID:G747674
C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate dei
C:Keywords: oxidoreductase
F:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 60.0%; Score 33; DB 2; Length 468;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
|||::||
Db 259 NKGTKWTSQ 268
RESULT 36
184555
phosphoglucuronate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Escherichia coli (strai
C:Species: Escherichia coli
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 11-Jun-1999
C:Accession: 184555
R:dykhuizen, D.E.; Green, L.
J. Bacteriol. 173, 7257-7268, 1991
A:Title: Recombination in Escherichia coli and definition of biological species.
A:Reference number: 141249; MUID:92041624; PMID:1938920
A:Accession: 184555
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-468 <RES>
C:Cross-references: GB:M64325; NID:G146929; PIDN:AAA24203.1; PID:G146930
C:Superfamily: phosphoglucuronate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate dei
C:Keywords: oxidoreductase
F:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <HIB>
Query Match 60.0%; Score 33; DB 2; Length 468;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 NKGTQOYTDQ 10
|||::||
Db 259 NKGTKWTSQ 268

RESULT 37

141249

phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Escherichia coli

C:Species: Escherichia coli

C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 11-Jun-1999

C:Accession: 141249

R: Dykhuizen, D.E.; Green, L.

J. Bacteriol. 173, 7257-7268, 1991

A:Title: Recombination in Escherichia coli and definition of biological species.

A:Reference number: 141249; PMID:92041624; PMID:1938920

A:Accession: 141249

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-468 <RES>

A:Cross-references: GB:M64324; NID:g146245; PIDN:AAA23925.1; PID:g146246

C:Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

C:Keywords: oxidoreductase

P:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <H1B>

Query Match

Best Local Similarity 60.0%; Score 33; DB 2; Length 468;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTDDQ 10

DB 259 NKGTGKMTSQ 268

RESULT 38

141250

phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44) - Escherichia coli

C:Species: Escherichia coli

C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 11-Jun-1999

C:Accession: 141250

R: Dykhuizen, D.; Bronner, D.; MacLachlan, P.R.; Dodgson, C.; Kido, N.; Whitfield, C.

J. Bacteriol. 176, 3126-3139, 1994

A:Title: Cloning and analysis of duplicated rfbM and rfbK genes involved in the formation of capsular polysaccharide.

A:Reference number: 141250; PMID:94252978; PMID:7515042

A:Accession: 141250

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-468 <RES>

A:Cross-references: GB:L27646; NID:g441134; PIDN:AAA21136.1; PID:g441135

C:Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

C:Keywords: oxidoreductase

P:6-285/Domain: 3-hydroxyisobutyrate dehydrogenase homology <H1B>

Query Match

Best Local Similarity 60.0%; Score 33; DB 2; Length 468;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTDDQ 10

DB 259 NKGTGKMTSQ 268

RESULT 39

F90982

gluconate-6-phosphate dehydrogenase [imported] - Escherichia coli (strain O157:H7, subsp

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: F90982

R: Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; PMID:21156231; PMID:11258796

A:Accession: F90982

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA836253.1; PID:g13362298; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC92830
 C:Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

Query Match

Best Local Similarity 60.0%; Score 33; DB 2; Length 468;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTDDQ 10

DB 259 NKGTGKMTSQ 268

RESULT 40

D85828

gluconate-6-phosphate dehydrogenase [imported] - Escherichia coli (strain O157:H7, subsp

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001

C:Accession: D85828

R: Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Hiller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; PMID:21074935; PMID:11206551

A:Accession: D85828

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 <STO>

A:Cross-references: GB:AE005174; NID:g12516213; PIDN:ANG57088.1; GSPDB:GN00145; UWGP:Z31

A:Experimental source: strain O157:H7, substrain EDL933

C:Genes: gnd

C:Superfamily: phosphogluconate dehydrogenase (decarboxylating); 3-hydroxyisobutyrate de

Query Match

Best Local Similarity 60.0%; Score 33; DB 2; Length 468;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTDDQ 10

DB 259 NKGTGKMTSQ 268

Search completed: August 20, 2003, 09:32:38
 Job time : 12.2371 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 5.7732 Seconds
(without alignments)
81.457 Million cell updates/sec

Title: US-09-991-433-5

Perfect score: 55

Sequence: 1 NKGTOQYTDQ 10

Scoring table: BLOSUM62
Gapex 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	781	1 COAT_PAVHB	P07299 human parvo
2	34	61.8	546	1 CH60_ACTAC	P46398 actinobacill
3	34	61.8	546	1 CH60_ACTPL	P94166 actinobacill
4	34	61.8	547	1 CH60_PASWU	O59687 pasteurilla
5	34	61.8	548	1 CH60_HABIN	P43733 haemophilus
6	34	61.8	652	1 C3BA_BACTU	O06117 bacillus th
7	34	61.8	659	1 C3BA_BACTO	P17969 bacillus th
8	34	61.8	872	1 CLPB_SYNY3	P74361 synchocyst
9	33	60.0	224	1 GSHU_CABEL	O95003 caenorhabdi
10	33	60.0	238	1 TRY3_SALSA	P35033 salmo salar
11	33	60.0	244	1 TRY2_XENLA	P70059 xenopus lae
12	33	60.0	246	1 UBIG_XYTRA	O09685 xylorella f
13	33	60.0	310	1 IFRH_TOBAC	P52579 nicotiana t
14	33	60.0	424	1 CRT3_ARATH	O04153 arabidopsis
15	33	60.0	445	1 6PGD_CITDM	P41581 citrobacter
16	33	60.0	445	1 6PGD_CITDM	P41582 citrobacter
17	33	60.0	445	1 6PGD_CITDM	P41583 citrobacter
18	33	60.0	445	1 6PGD_CITDM	P41584 escherichia
19	33	60.0	445	1 6PGD_KLEBL	P41575 klebsiella
20	33	60.0	445	1 6PGD_KLEBL	P41576 klebsiella
21	33	60.0	445	1 6PGD_SHIBO	P41577 shigella bo
22	33	60.0	445	1 6PGD_SHIBO	P41578 shigella dy
23	33	60.0	445	1 6PGD_SHIBO	P41579 shigella dy
24	33	60.0	445	1 6PGD_SHIBO	P41580 shigella so
25	33	60.0	468	1 6PGD_ECOLI	P33754 escherichia
26	33	60.0	468	1 6PGD_ECOLI	P00350 escherichia
27	33	60.0	468	1 6PGD_ECOLI	P41576 klebsiella
28	33	60.0	468	1 6PGD_SALTY	P14062 salmonella
29	33	60.0	468	1 6PGD_SALTY	P37756 shigella fl
30	33	60.0	1037	1 YHIV_ECOLI	P37673 escherichia
31	33	58.2	2352	1 MOKC_SCHPO	O9uul4 schizosach
32	33	58.2	181	1 Y896_VIBVU	O8ddr1 vibrio vuln
33	33	58.2	241	1 TR11_GADMO	P16049 gadus mornu
34	33	58.2	255	1 UPPS_METTH	O26334 mechanobact

34	32	58.2	262	1 KKH1_LACIA	Q02149 lactococcus
35	32	58.2	290	1 MREC_BACSU	Q01466 bacillus su
36	32	58.2	487	1 YOW5_CABEL	P06561 caenorhabdi
37	32	58.2	1063	1 DPOW_CLAPU	P22373 claviopsis p
38	32	58.2	1164	1 PHVD_ARATH	P42497 arabidopsis p
39	31	56.4	200	1 PHNH_RHIME	Q52985 rhizobium m
40	31	56.4	204	1 YPOL_ACICA	P07778 actinobact
41	31	56.4	228	1 BIOD_VIBPA	O87qn3 vibrio para
42	31	56.4	288	1 SSRA_ONCMY	P45433 oncorhynch
43	31	56.4	330	1 PHLC_STAPU	P09978 staphylococ
44	31	56.4	360	1 H1S8_BACSU	P17731 bacillus su
45	31	56.4	424	1 HD43_MOUSE	O88895 mus musculus

ALIGNMENTS

RESULT 1	ID	COAT_PAVHB	STANDARD	PRT	781 AA.
AC	P07259	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)				
DT	01-APR-1990 (Rel. 14, Last annotation update)				
DE	Probable coat protein VPI.				
OS	Human parvovirus B19.				
OC	Virusess: ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.				
OX	NCBI_TaxID=10798;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Isolate AU;				
RX	MEDLINE=86200451; PubMed=3701931;				
RA	Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;				
RT	"Nucleotide sequence and genome organization of human parvovirus B19 isolated from the serum of a child during aplastic crisis."				
RL	J. Virol. 58:921-936 (1986).				
CC	-1- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announce/ or send an email to license@ebi.ac.uk).				
CC	EMBL: M13178; AAA6667.1; -.				
DR	PIR: A24299; VCPV19.				
DR	InterPro: IPR001403; Parvo coat.				
DR	Pfam: PF00740; Parvo coat; 1.				
KW	Coat protein; Glycoprotein.				
FT	CARBOHYD 46				
FT	CARBOHYD 184				
FT	CARBOHYD 220				
FT	CARBOHYD 293				
FT	SEQUENCE 781 AA; 86015 MW; 866254DBD0576B07 CRC64;				
QY	1 NKGTOQYTDQ 10				
DB	622 NKGTOQYTDQ 631				
Query Match	100.0%;				
Best Local Similarity	100.0%;				
Matches	10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
RESULT 2					
ID	CH60_ACTAC	STANDARD	PRT	546 AA.	
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				

```

DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR GROEL OR MOP.
OS Actinobacillus actinomycetecombitans (Haemophilus
OC actinomycetecombitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-39.
RC STRAIN=Y4;
RX MEDLINE=96017061; PubMed=7567064;
RA Nakano T., Inai Y., Yamawata Y., Kuszaki-Nagata T., Nagaoka S.,
RA Okanashi N., Koga T., Nishihara T.,
RT "Molecular and immunological characterization of a 64-kDa protein of
RT Actinobacillus actinomycetecombitans";
RL Oral Microbiol. Immunol. 10:151-159(1995).
CC
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D28817; BAA05977.1; -
DR HSSP: P06139; IGRL.
DR HAMAP: MF_00600; -; 1.
DR InterPro: IPR001844; Chaperin_Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
KM
FT INIT MET 0
SQ SEQUENCE 546 AA; 57312 MW; 6249E0B46334F66 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 546;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 NKGTQOYTD 9
| | | | |
478 NAGTEQYGD 486

DB

```

```

RT pleuropneumoniae."
RL FEMS Microbiol. Lett. 147:11-16(1997).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U55016; AAB51437.1; -
DR HSSP: P06139; IGRL.
DR HAMAP: MF_00600; -; 1.
DR InterPro: IPR001844; Chaperin_Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KW Chaperone; ATP-binding.
KM
FT INIT MET 0
SQ SEQUENCE 546 AA; 57513 MW; B84BB72C9BD3DB56 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 546;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 NKGTQOYTD 9
| | | | |
479 NAGTEQYGD 487

DB

```

```

RESULT 3
CH60_ACTPL STANDARD; PRT; 546 AA.
AC P94166;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR GROEL OR MOP.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4074 / Serotype 1;
RX MEDLINE=97189570; PubMed=9037757;
RA Vezina G., Strole M., Claitoux N., Botsinot M.;
RT "Cloning and characterization of the groEL locus from Actinobacillus

```

```

RESULT 4
CH60_PASNU STANDARD; PRT; 547 AA.
AC Q59687;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR GROEL OR MOP OR Pm1107.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P1059;
RX MEDLINE=96105224; PubMed=8529887;
RA Love B.C., Hansen L.M., Hirsch D.C.;
RT "Cloning and sequence of the groEL heat-shock operon of Pasteurella
RT multocida."
RL Gene 166:179-180(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=96114586; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).

```

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30165; AAA84916.1; -.
DR EMBL; AB006151; AAK03191.1; -.
DR PIR; J04519; J04519.
DR HSSP; P06139; 1JON.
DR HAMAP; MF_00600; -; 1.
DR InterPro; IPR001844; Chaperin Cpn60.
DR InterPro; IPR002423; Cpn60/TCF-1.
DR Pfam; PF00118; Cpn60_TCF1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCPL.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
DR Chaperone; ATP-binding; Complete proteome.
KW CONFLICT 424 424 S -> N (IN REF. 1).
FT CONFLICT 464 464 I -> V (IN REF. 1).
SQ SEQUENCE 547 AA; 57291 MW; EBB95B9F2358B55E CRC64;

Query Match 61.8%; Score 34; DB 1; Length 547;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQGYTD 9
Db 479 NAGTEQYGD 487

RESULT 5
CH60_HA6IN STANDARD; PRT; 548 AA.
AC P43733;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN GROEL OR MOGA OR GROEL OR HI0543.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
NCBI_TaxID=727;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RA Rd.";
RL Science 269:496-512(1995).

-1- FUNCTION: Prevents misfolding and promotes the refolding and
-1- proper assembly of unfolded polypeptides generated under stress
-1- conditions (By similarity).
-1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
-1- 7 subunits (By similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32736; AAC22201.1; -.
DR PIR; G64076; G64076.
DR HSSP; P06139; 1JON.
DR TIGR; H10543; -.
DR HAMAP; MF_00600; -; 1.
DR InterPro; IPR001844; Chaperin Cpn60.
DR InterPro; IPR002423; Cpn60/TCF-1.
DR Pfam; PF00118; Cpn60_TCF1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCPL.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
DR Chaperone; ATP-binding; Complete proteome.
SQ SEQUENCE 548 AA; 57577 MW; C44066AAC1B62159 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 548;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQGYTD 9
Db 480 NAGTEQYGD 488

RESULT 6
C3BB_BACTU STANDARD; PRT; 652 AA.
AC Q06117; Q45717;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Peptidicidal crystal protein cry3bb (Insecticidal delta-endotoxin
DE CRYIIB(b)) (Crystalline entomocidal protoxin) (74 kDa crystal
DE protein).
GN CRY3BB OR CRYIIB(B) OR CRYIIB2.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=EG4961;
RX MEDLINE=93119147; PubMed=1476436;
RA Donovan W.P., Rupar M.J., Slaney A.C., Malvar T., Gawron-Burke M.C.,
RA Johnson T.B.;
RA "Characterization of two genes encoding Bacillus thuringiensis
RA insecticidal crystal proteins toxic to Coleoptera species.";
RL Appl. Environ. Microbiol. 58:3921-3927(1992).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=NRL B-1655 / EG5144;
RA Donovan W.P., Rupar M.J., Slaney A.C.;
RA "Bacillus thuringiensis cryIIIC, (b) protein toxic to coleopteran
RA insects.";
RL Patent number US5378625, 03-JAN-1995.

-1- FUNCTION: PROMOTES COLICIDIOSMOTIC LYSIS BY BINDING TO THE MIDGUT
-1- EPITHELIAL CELLS OF COLEOPTERA. HAS MODERATE LEVEL OF TOXICITY TO
-1- SOUTHERN CORN ROOTWORM.
-1- SUBUNIT: Monomer.
-1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
-1- SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART OF
-1- THE SPORE COAT.
-1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-
-1- TERMINUS.
-1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M89794; AAA2234.1; -
DR EMBL; U31633; AAA74198.1; -
DR PIR; I39811; I39811.
DR PDB; 1J16; 19-SRP-01.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation; 3D-structure.
KM VARIANT 21 21 Q -> P (IN STRAIN EG5144).
FT VARIANT 97 97 N -> D (IN STRAIN EG5144).
FT VARIANT 289 289 I -> V (IN STRAIN EG5144).
FT VARIANT 352 352 F -> S (IN STRAIN EG5144).
FT VARIANT 417 419 VHL -> IYF (IN STRAIN EG5144).
FT VARIANT 451 451 S -> G (IN STRAIN EG5144).
FT VARIANT 590 590 L -> I (IN STRAIN EG5144).
FT VARIANT 600 600 K -> I (IN STRAIN EG5144).
FT VARIANT 624 624 K -> T (IN STRAIN EG5144).
SQ SEQUENCE 652 AA; 74385 MW; 63048332CDE8CC6 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 652;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KGTQOYTD 9
Db 235 KLTQOYTD 242

RESULT 7
C3BA_BACTO STANDARD; PRT; 659 AA.
AC P17959;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Peptidic acid crystal protein cry3Ba (insecticidal delta-endotoxin
DE CryIIIBa) (Crystalline entomocidal protoxin) (75 kDa crystal
DE protein).
CRY3BA OR CRYIIIB(A) OR CRYIIIB.
Bacillus thuringiensis (subsp. tolworthi).
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_Taxid=1442;

RT Nucleic Acids Res. 18:1305-1305(1990).
RT isolate of Bacillus thuringiensis subsp. tolworthi.";
RT FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPIITHELIAL CELLS OF COLEOPTERA.
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.
CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X17123; CAA34983.1; -
DR EMBL; A07234; CAA00645.1; -
DR PIR; S10228; S10228.
DR HSSP; P07130; 1DLG.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation.
SQ SEQUENCE 659 AA; 75159 MW; 5A5B214FP84168CA CRC64;

Query Match 61.8%; Score 34; DB 1; Length 659;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KGTQOYTD 9
Db 243 KLTQOYTD 250

RESULT 8
CLPB_SYNY3 STANDARD; PRT; 872 AA.
AC P74361;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CLPB protein.
CN CLPB OR SLR1641.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_Taxid=1148;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kikura T.,
RA Hosouchi T., Matsumoto A., Muraki A., Nakazaki N., Nario K.,
RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II..Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- SUBUNIT: THOUGHT TO BE AN ATPASE SUBUNIT OF AN INTRACELLULAR
CC ATP-DEPENDENT PROTEASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CLPB/CLPB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D09914; BAA18456.1; -
DR PIR; S76197; S76197.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003599; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_c1pA/B.
DR InterPro; IPR004176; Clp_N.
DR Pfam; PF00004; AAA; 2.
DR Pfam; PF02861; Clp_N; 2.
DR PRINTS; PR00300; CLPPTASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPB_1; 1.

DR PROSITE; PS00871; CLPAB_2; 1.
KW Chaperone; ATP-binding; Repeat; Complete proteome.
FT DOMAIN 163 411 I.
FT NP_BIND 537 728 II.
FT NP_BIND 208 215 ATP (POTENTIAL).
FT NP_BIND 611 618 ATP (POTENTIAL).
SQ SEQUENCE 872 AA; 98122 MW; 7EAA8486C7D8D4D5 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 872;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 KGTQGYTDQ 10
DB 146 RGTQGYTDQ 154

RESULT 9

CC GSHU CAEEL STANDARD; PRT; 224 AA.
OY5003;

DR 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutathione peroxidase precursor (EC 1.11.1.9).
GN C11B4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodetidae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Mortimore B.;

RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 2 glutathione + H(2)O(2) = oxidized
CC glutathione + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -1- SIMILARITY: Belongs to the glutathione peroxidase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; Z81015; CAB02655.1; -
PIR; T19190; T19190.

DR HSSP; P00435; IGPI.
DR WormPep; C11B4.2; CE08102.
DR InterPro; IPR000889; Glut_peroxidase.
DR Pfam; PF00255; GSHPx; 1.
DR PRINTS; PR01011; GLUTPROXDASE.
DR PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
DR PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
KW Oxidoreductase; Peroxidase; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 224 GLUTATHIONE_PEROXIDASE.
FT ACT_SITE 73 73 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC... (POTENTIAL).
SQ SEQUENCE 224 AA; 25556 MW; F02D055246EDE2F1 CRC64;

Query Match 60.0%; Score 33; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TQGYTD 9
DB 76 TQGYTD 81

RESULT 10

TRY2_XENLA STANDARD; PRT; 238 AA.
AC P35033.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Trypsin III precursor (EC 3.4.21.4) (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_Taxid=8030;
RN (1)
RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;
RX MEDLINE=96035908; PubMed=7556223;
RA Male R., Lorens J.B., Smale A.O., Torrisen K.R.;
RT "Molecular cloning and characterization of anionic and cationic
RT variants of trypsin from Atlantic salmon.";
RL Eur. J. Biochem. 232:677-685(1995).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO TRYPSIN FAMILY S1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; X70074; CAA49679.1; -
PIR; S6657; S31779.
DR PDB; 1A0J; 13-JAN-99.

DR MEROPS; S01.151; -
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_Ser; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family; 3D-structure.

FT NON_TER 1 1 POTENTIAL.
FT SIGNAL <1 7 ACTIVATION PEPTIDE.
FT PROPEP 8 15 TRYPsin III.
FT CHAIN 16 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 55 55 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 99 99 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 192 192 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 22 152 BY SIMILARITY.
FT DISULFID 40 56 BY SIMILARITY.
FT DISULFID 124 225 BY SIMILARITY.
FT DISULFID 131 198 BY SIMILARITY.
FT DISULFID 163 217 BY SIMILARITY.
FT DISULFID 188 212 BY SIMILARITY.
FT SITE 186 186 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 238 AA; 25389 MW; AE799B80E8393023 CRC64;

Query Match 60.0%; Score 33; DB 1; Length 236;
Best Local Similarity 55.6%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NKGTQGYTD 9
DB 73 NKGTQGYTD 81

RESULT 11
TRY2_XENLA STANDARD; PRT; 244 AA.

AC P70059; 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin precursor (EC 3.4.21.4).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang K., Lytle L., Gan L., Hood L.E.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U72330; AAB1274.1; -.
 CC HSSP: P00763; IDPO.
 DR MEROPS: S01.258;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PRO0722; CHYMOTRYPsin.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolyase; Serine protease; Digestion; Zymogen; Signal;
 KW Multigene family.
 FT SIGNAL 1 15 BY SIMILARITY.
 FT PROPEP 16 21 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 22 244 TRYPsin.
 FT ACT_SITE 61 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 198 198 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 28 158 BY SIMILARITY.
 FT DISULFID 46 62 BY SIMILARITY.
 FT DISULFID 130 231 BY SIMILARITY.
 FT DISULFID 137 204 BY SIMILARITY.
 FT DISULFID 169 183 BY SIMILARITY.
 FT DISULFID 194 218 BY SIMILARITY.
 FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 244 AA; 26079 MW; C63729CB3300B23 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 244;
 Best Local Similarity 55.6%; Pred. No. 24;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 KNGTQGYTD 9
 DB 79 NKGTEGFID 87
 RESULT 12
 ID UBIG_XYLPA STANDARD; PRT; 246 AA.
 AC Q9PAM5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
 DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMS
 DE methyltransferase).

GN UBIG OR XP2471.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 NC NCB1_Taxid=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=9ASC;
 RC MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britton M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Doroty H.,
 RA Facinanci A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Fulton L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnsbeil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos B.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
 RA Petxoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pasquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silveira M.L.Z., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldans J., Zetbal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
 RL Nature 406:151-159(2000).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubi9quinone-9.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB004055; AAP85269.1; -.
 DR PIR: H82553; H82553.
 DR HAMAP: MF_00472; -; 1.
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 246 AA; 27121 MW; 5212107D53633D3P CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 246;
 Best Local Similarity 75.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 KGTQGYTD 9
 DB 182 KGTQGYKD 189
 RESULT 13
 ID IFRH_TOBAC STANDARD; PRT; 310 AA.
 AC P52579;
 DT 01-OCT-1996 (Rel. 34, Created)

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Isoflavone reductase homolog A622 (EC 1.3.1.-).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
OK NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Burley 21; TISSUE=Root;
RX MEDLINE=94312878; PubMed=8038607;
RA Hibi N., Higashiguchi S., Hashimoto T., Yamada Y.;
RT "Gene expression in tobacco low-nicotine mutants.";
RL Plant Cell 6:723-735(1994).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: STRONG, TO ISOFLAVONE REDUCTASE.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D28505; BAA05866.1; -
DR PIR; T02202; T02202.
DR InterPro; IPR003866; Isoflav_reduct.
DR Pfam; PF02716; Isoflavone_redu.1.
RW Oxidoreductase; NADP.
SQ SEQUENCE 310 AA; 34654 MW; 03A7F3BB6D451885 CRC64;

Query Match 60.0%; Score 33; DB 1; Length 310;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GTGQYTDQ 10
Db 87 GGQGFQTD 94

RESULT 14
CRTL ARATH STANDARD; PRT; 424 AA.
AC 004153; Q8SJE7;
DE 16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
GN Calreticulin 3 precursor.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OK NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97303616; PubMed=9159940;
RA Nelson D.E., Glaunsinger B., Bohmert H.U.;
RT "Abundant accumulation of the calcium-binding molecular chaperone
calreticulin in specific floral tissues of Arabidopsis thaliana.";
RL Plant Physiol. 114:29-37(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altabi H., Araujo R., Bowman C.I., Brooks S.Y.,
RA Buehler E., Conn A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,

```

```

RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lurce J.S., Maiti R., Marzilli A.,
RA Miltschev J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T.,
RA Pat G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana."
RL Nature 408:816-820(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U66345; AAC49697.1; -
DR EMBL; AC006932; AAF22902.1; AUT SEQ.
DR InterPro; IPR001580; Calreticulin.
DR InterPro; IPR000886; ER_target.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN 1; 1.
DR PROSITE; PS00804; CALRETICULIN 2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; FALSE NEG.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein;
KW Multigene family.
FT SIGNAL 1..28
FT CHAIN 29..424
FT CARBOHYD 97..97
FT SITE 421..424
FT CONFLICT 279..279
SQ SEQUENCE 424 AA; 49904 MW; 650E0AE8342F0B97 CRC64;

Query Match 60.0%; Score 33; DB 1; Length 424;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NKGTQYTD 9
Db 72 NKGIQTYND 80

RESULT 15
6PGD CITAM STANDARD; PRT; 445 AA.
AC 004158;
DE 01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
(Dragment).
GN GND.
OS Citrobacter amalonaticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OK NCBI_TaxID=35703;
RN [1]
RP SEQUENCE FROM N.A.

```

CC STRAIN-CT28;
 RX MEDLINE=95024018; PubMed=7937867;
 RA Nelson K., Selandar R.K.;
 RT "intergeneric transfer and recombination of the 6-phosphogluconate
 dehydrogenase gene (gnd) in enteric bacteria";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucuronate + NADP(+) = D-ribulose
 5-phosphate + CO(2) + NADPH.
 CC -1- PATHWAY: Hexose monophosphate shunt.
 CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL, U14426; AAC43773.1; -
 CC PIR, I40629; I40629.
 CC HSSP, P00349; 2PGD.
 CC InterPro, IPR006183; 6PGD.
 CC InterPro, IPR006114; 6PGD_C.
 CC InterPro, IPR006113; 6PGD_decarbox.
 CC InterPro, IPR006115; 6PGD_NAD.
 CC InterPro, IPR006184; 6PGDom.
 CC Pfam, PF00393; 6PGD; 1.
 CC Pfam, PF03446; NAD_binding_2; 1.
 CC PRINTS; PR00076; 6PGDHDRGNASE.
 CC TIGRFAMs; TIGR00873; gnd; 1.
 CC PROSITE; PS00461; 6PGD; 1.
 CC Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP.
 FT NON_TER 1 1
 FT VARIANT 445 445
 FT NON_TER 445 445
 FT NON_TER 445 445
 SQ SEQUENCE 445 AA; 48887 MW; A3771DFC9678EEF1 CRC64;

 CC Query Match 60.0%; Score 33; DB 1; Length 445;
 CC Best Local Similarity 60.0%; Pred. No. 47;
 CC Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 CC 1 NKGTQGYTDQ 10
 CC |||||::||
 CC Db 248 NKGTGKWTQ 257

CC STRAIN-CT28;
 RX MEDLINE=95024018; PubMed=7937867;
 RA Nelson K., Selandar R.K.;
 RT "intergeneric transfer and recombination of the 6-phosphogluconate
 dehydrogenase gene (gnd) in enteric bacteria";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucuronate + NADP(+) = D-ribulose
 5-phosphate + CO(2) + NADPH.
 CC -1- PATHWAY: Hexose monophosphate shunt.
 CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL, U14424; AAC43774.1; -
 CC EMBL, U14425; AAC43775.1; -
 CC EMBL, U14427; AAC43776.1; -
 CC EMBL, U14428; AAC43777.1; -
 CC EMBL, U14429; AAC43778.1; -
 CC EMBL, U14432; AAC43779.1; -
 CC PIR, I40681; I40681.
 CC PIR, I40682; I40682.
 CC PIR, I40684; I40684.
 CC PIR, I40685; I40685.
 CC HSSP, P00349; 2PGD.
 CC InterPro, IPR006183; 6PGD.
 CC InterPro, IPR006114; 6PGD_C.
 CC InterPro, IPR006113; 6PGD_decarbox.
 CC InterPro, IPR006115; 6PGD_NAD.
 CC InterPro, IPR006184; 6PGDom.
 CC Pfam, PF00393; 6PGD; 1.
 CC Pfam, PF03446; NAD_binding_2; 1.
 CC PRINTS; PR00076; 6PGDHDRGNASE.
 CC TIGRFAMs; TIGR00873; gnd; 1.
 CC PROSITE; PS00461; 6PGD; 1.
 CC Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP.
 FT NON_TER 1 1
 FT VARIANT 31 31 V -> G (IN STRAIN CT45).
 FT VARIANT 206 206 E -> D (IN STRAIN CT42).
 FT VARIANT 300 300 A -> E (IN STRAIN CT27).
 FT NON_TER 445 445
 FT NON_TER 445 445
 SQ SEQUENCE 445 AA; 48829 MW; 9DB96925EC40EDC4 CRC64;

 CC Query Match 60.0%; Score 33; DB 1; Length 445;
 CC Best Local Similarity 60.0%; Pred. No. 47;
 CC Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 CC 1 NKGTQGYTDQ 10
 CC |||||::||
 CC Db 248 NKGTGKWTQ 257

```
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U14466; AAC43814.1; -.
CC PIR: I40709; I40709.
CC DR HSSP: P00349; 2PGD.
CC DR InterPro: IPR006183; 6PGD.
CC DR InterPro: IPR006114; 6PGD_C.
CC DR InterPro: IPR006113; 6PGD_NAD.
CC DR InterPro: IPR006184; 6PGD_NAD.
CC Pfam: PF00393; 6PGD; 1.
CC Pfam: PF03446; NAD binding 2; 1.
CC PRINTS: PR00076; 6PGDHRGNASE.
CC TIGRFAMs: TIGR00873; gnd; 1.
CC PROSITE: PS00461; 6PGD; 1.
CC Glucanate utilization; Oxidoreductase; Pentose shunt; NADP.
CC NON_TER 1
CC FT NON_TER 1
CC SQ SEQUENCE 445 AA; 48832 MW; 717E56C3D6F3BFAC CRC64;

Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQGYTDQ 10
Db 248 NKGTGKWTQ 257

RESULT 18
6PGD_ESCVU STANDARD; PRT; 445 AA.
AC P41574;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
(Gragment).
GND.
Escherichia vulneris.
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33821;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Selander R.K.;
RT "Intergenic transfer and recombination of the 6-phosphogluconate
RT dehydrogenase gene (gnd) in enteric bacteria.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL: U14465; AAC43815.1; -.
CC DR HSSP: P00349; 2PGD.
CC DR InterPro: IPR006183; 6PGD.
CC DR InterPro: IPR006114; 6PGD_C.
CC DR InterPro: IPR006113; 6PGD_NAD.
CC DR InterPro: IPR006184; 6PGD_NAD.
CC Pfam: PF00393; 6PGD; 1.
CC Pfam: PF03446; NAD binding 2; 1.
CC PRINTS: PR00076; 6PGDHRGNASE.
CC TIGRFAMs: TIGR00873; gnd; 1.
CC DR PROSITE: PS00461; 6PGD; 1.
CC Glucanate utilization; Oxidoreductase; Pentose shunt; NADP.
CC NON_TER 1
CC FT NON_TER 1
CC SQ SEQUENCE 445 AA; 48806 MW; 0315B8126C2A45FE CRC64;

Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQGYTDQ 10
Db 248 NKGTGKWTQ 257

RESULT 19
6PGD_KLEPL STANDARD; PRT; 445 AA.
AC P41575;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
(Gragment).
GND.
Klebsiella planticola (Raoultella planticola).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Raoultella.
OX NCBI_TaxID=575;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33531;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Selander R.K.;
RT "Intergenic transfer and recombination of the 6-phosphogluconate
RT dehydrogenase gene (gnd) in enteric bacteria.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
CC 5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR TIGRFAM6; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP.
FT NON_TER 1 1
SQ SEQUENCE 445 AA; 48829 MW; 48E581E9B9290681 CRC64;
Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 NKGTQOYTDQ 10
Db 248 NKGTGKMTSQ 257
RESULT 20
6PGD_KLEETE STANDARD; PRT; 445 AA.
AC P41577;
RA 01-NOV-1995 (Rel. 32, Created)
RT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
OS (Fragment).
OC Klebsiella terrigena.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Raoultella.
OX NCBI_TaxID=577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33257;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Selander R.K.;
RT "Intergenic transfer and recombination of the 6-phosphogluconate
dehydrogenase gene (gnd) in enteric bacteria."
RT Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14473; AAC3819.1; -.
DR HSSP; P00349; 2PGD.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD_decarb.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGDm.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRCGNASE.
DR TIGRFAM6; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP.
FT NON_TER 1 1
SQ SEQUENCE 445 AA; 48834 MW; 2FBD0E46CDAB69A8 CRC64;
Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQ 10
Db 248 NKGTGKMTSQ 257
RESULT 21
6PGD_SHIBO STANDARD; PRT; 445 AA.
AC P41578;
RA 01-NOV-1995 (Rel. 32, Created)
RT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
OS (Fragment).
OC Shigella boydii.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=621;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 8700;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Selander R.K.;
RT "Intergenic transfer and recombination of the 6-phosphogluconate
dehydrogenase gene (gnd) in enteric bacteria."
RT Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14469; AAC3820.1; -.
DR HSSP; P00349; 2PGD.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD_decarb.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGDm.
DR Pfam; PF00393; 6PGD; 1.
DR Pfam; PF03446; NAD binding 2; 1.
DR PRINTS; PR00076; 6PGDHRCGNASE.
DR TIGRFAM6; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
KW Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP.
FT NON_TER 1 1
SQ SEQUENCE 445 AA; 48833 MW; B90F886E7AD0CB0B CRC64;
Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 NKGTQOYTDQ 10
Db 248 NKGTGKMTSQ 257
RESULT 22
6PGD_SHIDY STANDARD; PRT; 445 AA.
AC P41579;
RA 01-NOV-1995 (Rel. 32, Created)
RT 01-NOV-1995 (Rel. 32, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
DE (Fragment).
GN GND.
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=622;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13313;
RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Sclander R.K.;
RT "Intergenic transfer and recombination of the 6-phosphogluconate
dehydrogenase gene (gnd) in enteric bacteria.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231 (1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14467; AAC43821.1; -.
DR HSSP; P00349; 2PGD.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD_decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR Pfam; PF00393; 6PGD.1.
DR Pfam; PF03446; NAD binding_2; 1.
DR PRINTS; PR00076; 6PGDHGNASE.
DR TIGRPFAMS; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
DR Glucanate utilization; Oxidoreductase; Pentose shunt; NADP.
KW NON TER 1
FT 445
FT 445
SQ SEQUENCE 445 AA; 48862 MW; 04C8E50B7AC0116E CRC64;

Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTQTDQ 10
DB 248 NKGTGKWTSTQ 257

RESULT 23
6PGD_SHISO STANDARD; PRT; 445 AA.
ID _6PGD_SHISO
AC P41580;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
DE (Fragment).
GN GND.
OS Shigella sonnei.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=624;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29930;

```

```

RX MEDLINE=95024018; PubMed=7937867;
RA Nelson K., Sclander R.K.;
RT "Intergenic transfer and recombination of the 6-phosphogluconate
dehydrogenase gene (gnd) in enteric bacteria.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231 (1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14470; AAC43835.1; -.
DR EMBL; U14440; AAC43790.1; -.
DR HSSP; P00349; 2PGD.
DR InterPro; IPR006183; 6PGD.
DR InterPro; IPR006114; 6PGD_C.
DR InterPro; IPR006113; 6PGD_decarbox.
DR InterPro; IPR006115; 6PGD_NAD.
DR InterPro; IPR006184; 6PGDm.
DR Pfam; PF00393; 6PGD.1.
DR Pfam; PF03446; NAD binding_2; 1.
DR PRINTS; PR00076; 6PGDHGNASE.
DR TIGRPFAMS; TIGR00873; gnd; 1.
DR PROSITE; PS00461; 6PGD; 1.
DR Glucanate utilization; Oxidoreductase; Pentose shunt; NADP.
KW NON TER 1
FT 445
FT 445
SQ SEQUENCE 445 AA; 48824 MW; 07F8D753DBB0366E CRC64;

Query Match 60.0%; Score 33; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NKGTQQTQTDQ 10
DB 248 NKGTGKWTSTQ 257

RESULT 24
6PG9_ECOLI STANDARD; PRT; 468 AA.
ID _6PG9_ECOLI
AC P37754;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
GN GND.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O9:K30:Hi2 / E69;
RX MEDLINE=94252978; PubMed=7515042;
RA Jayaraman P., Bronner D., MacLachlan R.P., Dodgson C., Kido N.,
RA Whitefield C.;
RT "Cloning and analysis of duplicated rfbM and rfbK genes involved in
the formation of GDP-mannose in Escherichia coli O9:K30 and
participation of rfb genes in the synthesis of the group I K30
capsular polysaccharide.";
RL J. Bacteriol. 176:3126-3139 (1994).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
5-phosphate + CO(2) + NADPH.
CC -1- PATHWAY: Hexose monophosphate shunt.
CC -----

```

CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, L27646; AAA2136.1; -.
 CC PIR, I41250; I41250.
 CC HSSP, P00349; 2PGD.
 CC InterPro: IPR006183; 6PGD.
 CC InterPro: IPR006114; 6PGD_C.
 CC InterPro: IPR006113; 6PGD-decarbox.
 CC InterPro: IPR006115; 6PGD_NAD.
 CC InterPro: IPR006184; 6PGdom.
 CC Pfam; PF00393; 6PGD; 1.
 CC Pfam; PF03446; NAD binding 2; 1.
 CC PRINTS; PR00076; 6FGDHRGNASE.
 CC TIGRfam; TIGR00873; gnd; 1.
 CC PROSITE; PS00461; 6PGD; 1.
 CC Glucanate utilization; Oxidoreductase; Pentose shunt; NADP.
 CC KW
 CC SEQUENCE 468 AA; 51625 MW; C13D94CFD78BPF3A CRC64;
 SQ
 Qy Query Match 60.0%; Score 33; DB 1; Length 468;
 Db Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 NKGTQGYTDQ 10
 Db 259 NKGTGKWTSSQ 268
 RESULT 25
 6PGD_ECOLI STANDARD; PRT; 468 AA.
 AC P00350; P78080;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 6-phosphoglucanate dehydrogenase, decarboxylating (EC 1.1.1.44).
 GN GND OR B2029.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CC NCBI_Taxid=562;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=84237540; PubMed=6329905;
 CC Nasoff M.S., Baker H.V. II, Wolf R.E. Jr.;
 CC "DNA sequence of the Escherichia coli gene, gnd, for
 CC 6-phosphoglucanate dehydrogenase.";
 CC Gene 27:253-264 (1984).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12 / MG1655;
 CC MEDLINE=97426617; PubMed=9278503;
 CC Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 CC Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 CC Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 CC Mau B., Shao Y.;
 CC "The complete genome sequence of Escherichia coli K-12.";
 CC Science 277:1453-1474 (1997).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12;
 CC MEDLINE=97251358; PubMed=9097040;
 CC Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
 CC Isono K., Kasai H., Kimura S., Kitakawa M., Kikagawa M.,
 CC Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,

RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
 RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
 RA Takada J., Takemoto K., Wada C., Yamamoto Y., Horuchi T.;
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 40.1-50.0 min region on the linkage map.";
 RT DNA Res. 3:379-392 (1996).
 RN [4]
 RN SEQUENCE OF 1-125 FROM N.A.
 RX MEDLINE=89126937; PubMed=2464736;
 RA Miller R.D., Dykhuizen D.E., Hartl D.L.;
 RT "Fitness effects of a deletion mutation increasing transcription of
 RT the 6-phosphoglucanate dehydrogenase gene in Escherichia coli.";
 RL Mol. Biol. Evol. 5:691-703 (1988).
 RN [5]
 RN SEQUENCE OF 1-125 FROM N.A.
 RX MEDLINE=88086897; PubMed=3275621;
 RA Barcak G.J., Wolf R.E. Jr.;
 RT "Comparative nucleotide sequence analysis of growth-rate-regulated
 RT gnd alleles from natural isolates of Escherichia coli and from
 RT Salmonella typhimurium LT-2.";
 RL J. Bacteriol. 170:372-379 (1988).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucanate + NADP(+) = D-ribulose
 CC 5-phosphate + CO(2) + NADPH.
 CC -1- PATHWAY: Hexose monophosphate shunt.
 CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, K02072; AAA23918.1; -.
 CC EMBL, AE000294; AAC75090.1; -.
 CC EMBL, D90841; BAA15869.1; -.
 CC EMBL, M23181; AAA23924.1; -.
 CC EMBL, M18956; AAA23919.1; -.
 CC EMBL, M18957; AAA23920.1; -.
 CC EMBL, M18960; AAA23922.1; -.
 CC PIR, D64968; DEECGC.
 CC HSSP; P00349; 2PGD.
 CC ECODBASE; C042.6; 6TH EDITION.
 CC Ecogene; EG10411; gnd.
 CC InterPro: IPR006183; 6PGD.
 CC InterPro: IPR006114; 6PGD_C.
 CC InterPro: IPR006113; 6PGD-decarbox.
 CC InterPro: IPR006115; 6PGD_NAD.
 CC InterPro: IPR006184; 6PGdom.
 CC Pfam; PF00393; 6PGD; 1.
 CC Pfam; PF03446; NAD binding 2; 1.
 CC PRINTS; PR00076; 6FGDHRGNASE.
 CC TIGRfam; TIGR00873; gnd; 1.
 CC PROSITE; PS00461; 6PGD; 1.
 CC Glucanate utilization; Oxidoreductase; Pentose shunt; NADP;
 CC Complete proteome.
 CC KW
 CC CONFLICT 306 P -> R (IN REF. 1).
 CC SEQUENCE 468 AA; 51481 MW; 62A32C84DC56D86 CRC64;
 SQ
 Qy Query Match 60.0%; Score 33; DB 1; Length 468;
 Db Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 NKGTQGYTDQ 10
 Db 259 NKGTGKWTSSQ 268
 RESULT 26
 6PGD_KLEPN STANDARD; PRT; 468 AA.
 ID 6PGD_KLEPN

AC P41576; Q48461;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
 GN GND.
 OS Klebsiella pneumoniae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Klebsiella.
 CX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Chedid;
 RX MEDLINE=95204345; PubMed=7896702;
 RA Arkawa Y., Wacharotayanakun R., Nagatsuka T., Ito H., Kato N.,
 RA Ohta M.; Wacharotayanakun R., Nagatsuka T., Ito H., Kato N.,
 RT "Genomic organization of the Klebsiella pneumoniae cps region
 responsible for serotype K2 capsular polysaccharide synthesis in the
 virulent strain Chedid."
 RL J. Bacteriol. 177:1788-1796 (1995).
 [2]
 RP SEQUENCE OF 12-456 FROM N.A.
 RC STRAIN=CM 7380;
 RX MEDLINE=95024018; PubMed=7937867;
 RA Nelson K., Selandier R.K.;
 RT "Intergenic transfer and recombination of the 6-phosphogluconate
 dehydrogenase gene (gnd) in enteric bacteria."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231 (1994)
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
 5-phosphate + CO(2) + NADPH.
 CC -1- PATHWAY: Hexose monophosphate shunt.
 CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D1242; BAA04786.1; -;
 DR EMBL; U14471; AAC43817.1; -;
 DR PIR; D56146; D56146.
 DR HSSP; P00349; 2PGD.
 DR InterPro; IPR006183; 6PGD.
 DR InterPro; IPR006114; 6PGD_C.
 DR InterPro; IPR006113; 6PGD_decarbox.
 DR InterPro; IPR006184; 6PGD_NAD.
 DR Pfam; PF00393; 6PGD_1.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.
 DR TIGRFAMs; TIGR00873; gnd; 1.
 DR PROSITE; PS00461; 6PGD; 1.
 KW Glucanone utilization; Oxidoreductase; Pentose shunt; NADP.
 FT CONFLICT 316 316 G -> E (IN REF. 2).
 FT CONFLICT 421 421 V -> F (IN REF. 2).
 SQ SEQUENCE 468 AA; 51328 MW; FFLBSE7655FDC90 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 468;
 Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTOQYTDQ 10
 DB 259 NKGTKMTSQ 268
 RESULT 27
 6PGD_SALTY STANDARD; PRT; 468 AA.
 ID 6PGD_SALTY

AC P14062;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
 GN GND OR STM2081.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 CX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE=89364685; PubMed=2671649;
 RA Reeves P., Stevenson G.;
 RT "Cloning and nucleotide sequence of the Salmonella typhimurium LT2
 gnd gene and its homology with the corresponding sequence of
 Escherichia coli K12."
 RL Mol. Gen. Genet. 217:182-184 (1989).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RA Dykhuizen D.E., Green L.;
 RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Portolillo S., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2."
 RL Nature 413:852-856 (2001).
 [4]
 RP SEQUENCE OF 1-57 FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE=91260454; PubMed=1710759;
 RA Jiang X.M., Neal B., Santiago F., Lee S.J., Romana L.K., Reeves P.R.;
 RT "Structure and sequence of the rfb (O antigen) gene cluster of
 Salmonella serovar Typhimurium (strain LT2)."
 RL Mol. Microbiol. 5:695-713 (1991).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
 5-phosphate + CO(2) + NADPH.
 CC -1- PATHWAY: Hexose monophosphate shunt.
 CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X15651; CAA33677.1; -;
 DR EMBL; M64332; AAA27137.1; -;
 DR EMBL; AE008792; AAL20985.1; -;
 DR EMBL; X56793; CAA40131.1; -;
 DR PIR; S04397; S04397.
 DR HSSP; P00349; 2PGD.
 DR ScyGene; SG10146; gnd.
 DR InterPro; IPR006183; 6PGD.
 DR InterPro; IPR006114; 6PGD_C.
 DR InterPro; IPR006113; 6PGD_decarbox.
 DR InterPro; IPR006115; 6PGD_NAD.
 DR InterPro; IPR006184; 6PGD_NAD.
 DR Pfam; PF00393; 6PGD_1.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.

DR TIGRFAMs; TIGR00873; gnd; 1.
 DR PROSITE; PS00461; 6PGD; 1.
 DR Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP;
 KW Complete proteome.
 SQ SEQUENCE 468 AA; 51395 MW; D8EB53A2DAADF7 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 468;
 Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTQGYTDQ 10
 DB 259 NKGTGKMTSQ 268
 RESULT 28
 ID 6PGD_SHIFL STANDARD; PRT; 468 AA.
 AC P37756;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 GN 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PE577 / Serotype 2a;
 RC MEDLINE=94131953; PubMed=7507920;
 RC Morona R., Mavris M., Fallarino A., Manning P.A.;
 RT "Characterization of the rfc region of Shigella flexneri.";
 RL J. Bacteriol. 176:733-747(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RC MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [3]
 RP SEQUENCE OF 12-456 FROM N.A.
 RC STRAIN=ATCC 29903;
 RC MEDLINE=95024018; PubMed=7937867;
 RA Nelson K., Selandar R.K.;
 RT "Intergenic transfer and recombination of the 6-phosphogluconate
 RT dehydrogenase gene (gnd) in enteric bacteria.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10227-10231(1994).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-gluconate + NADP(+) = D-ribulose
 CC 5-phosphate + CO(2) + NADPH.
 CC -1- PATHWAY: Hexose monophosphate shunt.
 CC -1- SIMILARITY: BELONGS TO THE 6-PHOSPHOGLUCONATE DEHYDROGENASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X71970; CAA50761.1; -;
 CC EMBL; AE015225; AA043630.1; -;
 CC EMBL; U04468; AAC3834.1; -;
 CC HSSP; P00349; 2PGD.

DR InterPro; IPR006183; 6PGD.
 DR InterPro; IPR006114; 6PGD_C.
 DR InterPro; IPR006113; 6PGD_decarbox.
 DR InterPro; IPR006115; 6PGD_NAD.
 DR InterPro; IPR006184; 6PGDm.
 DR Pfam; PF00393; 6PGD; 1.
 DR Pfam; PF03446; NAD_binding_2; 1.
 DR PRINTS; PR00076; 6PGDHDRGNASE.
 DR TIGRFAMs; TIGR00873; gnd; 1.
 DR PROSITE; PS00461; 6PGD; 1.
 KW Glucuronate utilization; Oxidoreductase; Pentose shunt; NADP.
 SQ SEQUENCE 468 AA; 51344 MW; 2203A0D82120CD61 CRC64;
 Query Match 60.0%; Score 33; DB 1; Length 468;
 Best Local Similarity 60.0%; Pred. No. 49;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NKGTQGYTDQ 10
 DB 259 NKGTGKMTSQ 268
 RESULT 29
 ID YHIV_ECOLI STANDARD; PRT; 1037 AA.
 AC P37637;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yHIV.
 GN YHIV OR B3514.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RC MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 CC -1- FUNCTION: COULD BE A DRUG EFFLUX PUMP (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -1- SIMILARITY: BELONGS TO THE ACB/ACRD/ACRF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U00039; AB18490.1; -;
 CC EMBL; AE000427; AAC76539.1; -;
 CC PIR; S47734; S47734.
 CC Ecogen; EG12241; yHIV.
 DR InterPro; IPR001036; Acetylvin_res.
 DR InterPro; IPR004764; HAB1.
 DR Pfam; PF00873; ACR_tran; 1.
 DR PRINTS; PR00702; AGRIFLAVINRP.
 DR TIGRFAMs; TIGR00915; 2A0602; 1.
 KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
 KW Complete proteome.
 FT DOMAIN 1 9 CYTOPLASMIC (BY SIMILARITY).
 FT TRANSMEM 10 28 1 (BY SIMILARITY).
 FT DOMAIN 29 339 PERIPLASMIC (BY SIMILARITY).
 FT TRANSMEM 340 359 2 (BY SIMILARITY).
 FT DOMAIN 360 365 CYTOPLASMIC (BY SIMILARITY).
 FT TRANSMEM 366 385 3 (BY SIMILARITY).

```

FT DOMAIN 386 391 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 392 413 4 (BY SIMILARITY).
FT DOMAIN 414 441 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 442 460 5 (BY SIMILARITY).
FT DOMAIN 461 473 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 474 496 6 (BY SIMILARITY).
FT DOMAIN 497 536 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 537 555 7 (BY SIMILARITY).
FT DOMAIN 556 870 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 871 890 8 (BY SIMILARITY).
FT DOMAIN 891 896 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 897 916 9 (BY SIMILARITY).
FT DOMAIN 917 922 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 923 944 10 (BY SIMILARITY).
FT DOMAIN 945 972 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 973 991 11 (BY SIMILARITY).
FT DOMAIN 992 1004 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 1005 1027 12 (BY SIMILARITY).
FT DOMAIN 1028 1037 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1037 11517 MM; 3867502160F51724 CRC64;
SEQUENCE

Query Match 60.0%; Score 33; DB 1; Length 1037;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 KGTQOYTD 9
Db 520 KSTQHYTD 527

RESULT 30
MOKK_SCHPO STANDARD; PRT; 2352 AA.
ID Q9YU14; O133605;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell wall alpha-1,3-glucan synthase mok12 (EC 2.4.1.183).
GN MOK12 OR SPB32H6.13C OR P1011.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=972;
RA Katayama S., Dai H., Arellano M., Perez P., Toda T.;
RA "Fission yeast alpha-glucan synthase Mok1 localizes closely with actin
RA and play a role essential for cell morphogenesis and protein kinase C
RA function.",
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=972;
RC MEDLINE=20089027; PubMed=10620777;
RA Mechida M., Yamazaki S., Kunihiro S., Tanaka T., Kuehida N., Jimo K.,
RA Hattawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,
RA Sakai M., Aoki K., Ogura K., Kudoh Y., Kikuchi H., Zhang M.Q.,
RA Yanagida M.;
RA "A 38 kb segment containing the cdc2 gene from the left arm of fission
RA yeast chromosome II: sequence analysis and characterization of the
RA genomic DNA and cDNAs encoded on the segment.",
RA Yeast 16:71-80(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

```

```

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Mblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymptre B.,
RA Weltens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Punnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,
RA Galibert F., Aves J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rocher M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usero D., Barrett B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.",
RA Nature 415:871-880(2002)
RT Nature 415:871-880(2002)
CC -1- CATALYTIC ACTIVITY: UDP-glucose + (alpha-D-glucosyl-(1,3))(N) =
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AB018381; BA476558.1;
CC DR EMBL; AB004534; BA421388.1; ALT_INIT.
CC DR EMBL; AL5590971; CAC37503.1;
CC DR PIR; T43431; T43431.
CC DR GeneDB; SPB32H6.13C;
CC DR InterPro; IPR006047; Alpha_amy1_cat.
CC DR InterPro; IPR001296; Glyco_transf_1.
CC DR Pfam; PF00534; Alpha-amyase; 1.
CC DR Pfam; PF00534; Glycosyltransf_1; 1.
CC DR Cell wall; Transferase; Glycosyltransferase.
CC KW SEQUENCE 2352 AA; 26561 MW; 78ADP9C2F7140BBA CRC64;
SQ

Query Match 60.0%; Score 33; DB 1; Length 2352;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 GTQOYTD 9
Db 543 GTQYTD 549

RESULT 31
Y896_VIBVU STANDARD; PRT; 181 AA.
ID Y896_VIBVU
AC Q8DRL;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0241 protein V10896.
GN V10896.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_Taxid=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RA "Complete genome sequence of Vibrio vulnificus CMCP6.",

```

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the UPF0241 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE016800; AAC0939.1; -
 DR HAMAP; MF_01058; -1.
 DR InterPro; IPR007336; DUF414.
 DR Pfam; PF04220; DUF414; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 181 AA; 20568 MW; 1CF629CFFPA7ABF CRC64;
 Query Match 58.2%; Score 32; DB 1; Length 181;
 Best Local Similarity 55.6%; Pred. No. 28;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 KGTQQTDD 10
 DB 121 RGLQGYVDE 129
 RESULT 32
 TR1L_GADMO STANDARD; PRT; 241 AA.
 AC P16049; Q91040; Q92156;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin I precursor (EC 3.4.21.4).
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pyloic caeca;
 RX MEDLINE=89210867; PubMed=2707266;
 RA Aegleeson B., Fox J.W., Bjarnason J.B.;
 RT "Purification and characterization of trypsin from the polkilotherm
 RT Gadus morhua";
 RL Eur. J. Biochem. 180:85-94(1989).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X76886; CAA54214.1; -
 DR FIR; S39047; S39047.
 DR HSRP; P00763; IDPO.
 DR MEROPS; S01.151; -.

DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KM Multigene family.
 FT SIGNAL 1 13
 FT PROPEP 14 19
 FT CHAIN 20 241
 FT ACT_SITE 59 59
 FT ACT_SITE 103 103
 FT ACT_SITE 195 195
 FT DISULFID 26 155
 FT DISULFID 44 60
 FT DISULFID 128 228
 FT DISULFID 135 201
 FT DISULFID 166 180
 FT DISULFID 191 215
 FT SITE 189 189
 FT SITE 25 25
 FT CONFLICT 27 28
 FT CONFLICT 43 43
 FT CONFLICT 49 52
 SQ SEQUENCE 241 AA; 25941 MW; 44BC9A0106AD1A68 CRC64;
 Query Match 58.2%; Score 32; DB 1; Length 241;
 Best Local Similarity 71.4%; Pred. No. 38;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NKGTQY 7
 DB 77 NEGTEQY 83
 RESULT 33
 UPFS_METH STANDARD; PRT; 255 AA.
 ID UPFS_METH
 AC 026334;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) (UPP
 DE synthetase) (di-trans-poly-cis-decaprenylcistransferase) (Undecaprenyl
 DE di-phosphate synthase) (UDS).
 DB MTH232.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Baehrtraden R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lamm W., Pochler B., Qiu D.,
 RA Spadafora R., Vlcare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Carnes A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-T., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- FUNCTION: GENERATES UNDECAPRENYL PYROPHOSPHATE (UPP) FROM
 CC ISOPRENYL PYROPHOSPHATE (IPP) (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Di-trans-poly-cis-decaprenyl diphosphate +
 CC isopentenyl diphosphate = diphosphate + di-trans-poly-cis-
 CC undecaprenyl diphosphate.
 CC -1- SIMILARITY: BELONGS TO THE UPP SYNTHETASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000810; AAB84738.1; --
 DR PIR: H69128; H69128.
 DR InterPro: IPR001441; UPP_synth.
 DR Pfam: PF01255; UPP_synthetase; 1.
 DR ProDom: PD003461; UPP_synth; 1.
 DR TIGRFAMs: TIGR00055; UPPS; 1.
 DR PROSITE: PS01066; UPP_SYNTHETASE; 1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 255 AA; 29900 MW; 370C6D8F85FEB8D CRC64;

 QY 2 KGTQGYTDQ 10
 Db 139 KSTEQYSDR 147

 RESULT 34
 KKH LACLA STANDARD; PRT; 262 AA.
 ID KKH_LACLA
 AC Q02149;
 DT 01-JUL-1993 (Rel. 26; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Probable aminoglycoside 3'-phosphotransferase (EC 2.7.1.95) (Kanamycin
 DE kinase).
 GN YMDK OR L1211.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCDO 2118;
 RX MEDLINE=93015709; PubMed=1400209;
 RA DeJorne C., Ehrlich S.D., Renault P.;
 RA "Histidine biosynthesis genes in Lactococcus lactis subsp. lactis";
 J. Bacteriol. 174:6571-6579(1992).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Winkler P., Manger S., Taillon O., Malarme K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RA "The complete genome sequence of the lactic acid bacterium Lactococcus
 RA lactis ssp. lactis IL1403";
 Genome Res. 11:731-753(2001).
 RL -1- CATALYTIC ACTIVITY: ATP + kanamycin 3'-
 phosphate.
 -1- SIMILARITY: TO OTHER AMINOGLYCOSIDE PHOSPHOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U92974; AAB81906.1; --
 DR EMBL: AE006353; AAK05309.1; --
 DR PIR: C86776; C86776.
 DR PIR: H45734; H45734.

DR InterPro: IPR002575; APH.
 DR Pfam: PF01636; APH; 1.
 KW Hypothetical protein; Antibiotic resistance; Transferase; Kinase;
 KW ATP-binding; Complete proteome.
 FT ACT SITE 187 187 BY SIMILARITY.
 FT CONFLICT 164 164 I -> T (IN REF. 1).
 SQ SEQUENCE 262 AA; 30720 MW; 62661D98FAA15527 CRC64;

 QY 1 NKGTQGYTDQ 10
 Db 228 NIGTDQYDR 237

 RESULT 35
 MREC BACSU STANDARD; PRT; 290 AA.
 ID MREC_BACSU
 AC 001466;
 DT 01-APR-1993 (Rel. 25; Created)
 DT 01-APR-1993 (Rel. 25; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Rod shape-determining protein mrec.
 GN MREC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=93015732; PubMed=1400225;
 RA Varley A.W., Stewart G.C.;
 RA "The divIVA region of the Bacillus subtilis chromosome encodes
 RA homologs of Escherichia coli septum placement (mncD) and cell shape
 RA (mreBCD) determinants";
 J. Bacteriol. 174:6729-6742(1992).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93015731; PubMed=1400224;
 RA Levin P.A., Margolis P.S., Setlow P., Losick R., Sun D.;
 RA "Identification of Bacillus subtilis genes for septum placement and
 RA shape determination";
 J. Bacteriol. 174:6717-6728(1992).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berrero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boutsier L., Brans A., Braun M., Briganti S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Enlian K.D., Errington J., Fabret C., Ferrari E., Follmer D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gim S.Y., Glaser P., Goiteau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hult M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kutsayashi Y., Koester P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.W., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Melillo R.P., Mizuno M., Moesli D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viarelli A., Wambat R., Wedler E., Wedler H., Weitzenecker T.,
 RA Viarelli P., Wambat R., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis".
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: INVOLVED IN FORMATION OF THE ROD SHAPE OF THE CELL. MAY
 CC ALSO CONTRIBUTE TO REGULATION OF FORMATION OF PENICILLIN-BINDING
 CC PROTEINS (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M95582; AAA22606.1; -
 CC EMBL; M96343; AAA22398.1; -
 CC EMBL; Z99118; CAB4762.1; -
 CC PIR; C45240; C45240.
 CC PIR; E69660; E69660.
 CC Subtilast; BG10327; mrec.
 CC InterPro; IPR005223; Mrec.
 CC Pfam; PF04085; Mrec; 1.
 CC TIGRfam; TIGR00219; mrec; 1.
 CC Cell shape; Complete proteome.
 CC CONFLICT 143 143 K -> F (IN REF. 2).
 CC CONFLICT 176 176 T -> P (IN REF. 2).
 CC NCBI_TaxID=5111;
 CC SEQUENCE 290 AA; 32140 MW; B9984D5B8414A5C CRC64;
 SQ
 Query Match 58.2%; Score 32; DB 1; Length 290;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NKGTQO 6
 DB 134 NKGTQO 139
 RESULT 36
 YOM5 CAEEL STANDARD; PRT; 487 AA.
 AC P30651;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 16-OCT-2001 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 RT Hypothetical protein ZK643.5 in chromosome III.
 RT ZK643.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br101 N2.
 RX MEDLINE=92168156; PubMed=1538779;
 RA Sullivan J., Du Z., Thomas K., Wilson R., Hillier L., Staden R.,
 RA Halloran N., Green P., Thierry-Mieg J., Qiu L., Dear S., Coulson A.,
 RA Craxton M., Durbin R., Berke M., Metzstein M., Hawkins T.,
 RA Ainscough R., Waterston R.;
 RT "The C. elegans genome sequencing project: a beginning";
 RT Nature 356:37-41(1992).
 RL Nature 356:37-41(1992).
 RN
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 21126; CAA77474.2; -
 CC WormPep; ZK643.5; CE24732.
 DR WormPep; ZK643.5; CE24732.
 KW Hypothetical protein.
 SQ SEQUENCE 487 AA; 55241 MW; BD6FD02D7055013 CRC64;
 Query Match 58.2%; Score 32; DB 1; Length 487;
 Best Local Similarity 75.0%; Pred. No. 82;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 KGTQOYTD 9
 DB 38 KGTQOYTD 45
 RESULT 37
 DPOW CLAPU STANDARD; PRT; 1063 AA.
 ID DPOW CLAPU
 AC P2373;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Probable DNA polymerase (EC 2.7.7.7).
 OS Claviceps purpurea (Ergot fungus).
 OG Mitochondrion.
 OG Plasmid pCL1.
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Hypocryomycetidae; Hypocreales; Clavicipitaceae; Claviceps.
 OX NCBI_TaxID=5111;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N;
 RX MEDLINE=89364677; PubMed=2770691;
 RA Oeser B., Tudymski P.;
 RT "The linear mitochondrial plasmid pCL1 of the phytopathogenic fungus
 RT Claviceps purpurea may code for a DNA polymerase and an RNA
 RT polymerase";
 RT Mol. Gen. Genet. 217:132-140(1989).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -1- MISCELLANEOUS: THIS DNA POLYMERASE REQUIRES A PROTEIN AS A PRIMER
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. STRONG TO
 CC DNA POLYMERASE OF OTHER FUNGAL AND PLANTS MITOCHONDRIAL PLASMIDS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15648; -; NOT_ANNOTATED_CDS.
 DR InterPro; IPR006172; DNA_POL_B.
 DR InterPro; IPR004868; DNA_POL_B_2.
 DR Pfam; PF03175; DNA_POL_B_2; 1.
 DR SMART; SM00486; POLB; 1.
 DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Plasmid; Mitochondrion.
 SQ SEQUENCE 1063 AA; 122540 MW; 3F5621BF3B6E147C CRC64;
 Query Match 58.2%; Score 32; DB 1; Length 1063;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NKGTQOYTD 9
 DB 1 NKGTQOYTD 9

DB 909 NNGTLYTDD 917

RESULT 38

PHYD_ARATH STANDARD; PRT; 1164 AA.

AC P42497; 023472;

DT 01-NOV-1995 (Rel. 32, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phytochrome D

GN PHYD OR AY4616250 OR DL4165C.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosidia II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID:3702;

RN [1]

SEQUENCE FROM N.A.

STRAIN=cv. Landsberg erecta;

MEDLINE=94325466; PubMed=8049367;

Clack T., Mathews S., Sharrock R.A.;

"The phytochrome apoprotein family in Arabidopsis is encoded by five genes: the sequences and expression of PHYD and PHYE.";

Plant Mol. Biol. 25:413-427(1994).

[2]

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

MEDLINE=96121113; PubMed=9461215;

Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C., Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L., Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P., Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T., Terry N., Gielens J., Villarroel R., De Clercq R., van Montagu M., Lecharny A., Aubourg S., Gylt, Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P., Entian K.-D., Rieger M., Schaefer M., Funk B., Mueller-Auer S., Silvey M., James R., Monfort A., Pons A., Pulgomech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B., Hilbert H., Duesterhoeft A., Moores T., Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Ansoerge W., Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Wewes H.-W., Klosteman S., Schueller C., Chalwatzis N.;

"Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana.";

Nature 391:485-488(1998).

[3]

SEQUENCE FROM N.A.

STRAIN=cv. Columbia;

MEDLINE=20083488; PubMed=10617198;

Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M., Weischaelger M., de Simone V., Obermaier B., Macho R., Mueller M., Kreis M., Delseny M., Pulgomech P., Watson M., Schmidheini T., Reichert B., Portetalle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Van der Schueren J., Grymoprez B., Chuang Y.-O., Vandenbusche F., Breken M., Wellens I., Voet M., Baatiens I., Aert R., Delfor E., Weitzenegger T., Boche G., Ramepberger U., Hilbert H., Braun M., Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W., Moeliman P., Klein lankhorst R., Rose M., Hauf J., Koeltter P., Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buyschaert C., Gielens J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McElay K., Mayes R., Portetalle A., Rajadream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Granderath K., Danner D., Herzl A., Neumann S., Agirion A., Vitale D., Ligouri R., Piravandi E., Maassen O., Quileley F., Clabaud G., Mendenlein A., Felber R., Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,

RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,

RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,

RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,

RA Heijnen L., Schwarz S., Scholler P., Heber S., France P., Belke C.,

RA Frishman D., Haase D., Lemcke K., Wewes H.-W., Stocker S.,

RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,

RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spliegel L.,

RA Sekhon M., Murray J., Sheel P., Cordes M., Abu-Theideh J.,

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,

RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,

RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,

RA Kramer J., Fulton L., Wadley E., Dante M., Pepin K., Hillier L.,

RA Nelson J., Speith J., Ryan E., Andrews S., Geisel C., Layman D.,

RA Du H., Ali J., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,

RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,

RA Ma P., Zhong J., Preston R., Vil D., Shaker M., Matero A., Shan R.,

RA Swaby I.K., Shohdy N., Hasegawa A., Rodriguez M., Hoffman J., Tili S.,

RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodi M., Johnson A.,

RA Chen E., Marra M., Martensen R., McCombie W.R.;

"Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.";

Nature 402:769-777(1999).

-1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RUBISCO-BIPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN, PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.

-1- SUBUNIT: Homodimer.

-1- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.

-1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.

-1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.

-1- SIMILARITY: Contains 1 histidine kinase domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X76609; CAA54072.1; -;

DR EMBL; Z97340; CAB10404.1; -;

DR EMBL; AL161543; CAB78667.1; -;

DR PIR; B71429; B71429.

DR InterPro; IPR003594; ATPbind_ATPase.

DR InterPro; IPR003018; GAF.

DR InterPro; IPR003617; His_kinase.

DR InterPro; IPR005467; His_kinase.

DR InterPro; IPR000014; PAS_domain.

DR InterPro; IPR001294; Phytochrome.

DR Pfam; PF01590; GAF; 1.

DR Pfam; PF02516; HATPase_C; 1.

DR Pfam; PF00512; HSKA; 1.

DR Pfam; PF00989; PAS; 2.

DR Pfam; PF00360; phytochrome; 1.

DR PRINTS; PR01033; PHYTOCHROME.

DR SMART; SM00065; GAF; 1.

DR SMART; SM00387; HATPase_C; 1.

DR SMART; SM00388; HSKA; 1.

DR SMART; SM00091; PAS; 2.

DR TIGRfam; TIGR00229; sensory_box; 2.

DR PROSITE; PS50109; HIS_KIN; 1.

DR PROSITE; PS50112; PAS; 2.

DR PROSITE; PS50245; PHYTOCHROME_1; 1.

DR PROSITE; PS50046; PHYTOCHROME_2; 1.

KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;

KW Repeat; Multigene family.
 FT DOMAIN 656 727 PAS 1.
 FT DOMAIN 790 861 PAS 2.
 FT BINDING 938 1157 HISTIDINE KINASE.
 FT BINDING 360 360 CHROMOPHORE (BY SIMILARITY).
 FT CONFLICT 425 425 L -> F (IN REF. 1).
 SQ SEQUENCE 1164 AA; 129267 MW; BB7CE19C50ACBAB CRC64;

Query Match 58.2%; Score 32; DB 1; Length 1164;
 Best Local Similarity 75.0%; Pred. No. 2, le+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 Db 57 NKAIQOYT 64

RESULT 39
 PHNH RHIME STANDARD; PRT; 200 AA.
 052985;
 30-MAY-2000 (Rel. 39, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 DE PhnH protein.
 GN PHNH OR RB1452 OR SMB20760.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymb (megaplasmid 2).
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=1021;
 RC MEDLINE=21396508; PubMed=1181431;
 RX Finn T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
 RA Golding B., Puehler A.;
 "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
 fixing endosymbiont Sinorhizobium meliloti."
 Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 -1- FUNCTION: BELONGS TO AN OPERON INVOLVED IN UTILIZATION OF
 GLYPHOSATE AND OTHER PHOSPHONATES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; M96263; AAA26351.1; -
 DR EMBL; AL603647; CAC49852.1; -
 DR PIR; D96023; D96023.
 KW Alkylphosphonate uptake; Plasmid; Complete proteome.
 SQ SEQUENCE 200 AA; 21253 MW; 3C8E48F736F5AFC1 CRC64;

Query Match 56.4%; Score 31; DB 1; Length 200;
 Best Local Similarity 62.5%; Pred. No. 50;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GTOOYTDO 10
 Db 117 GTOEYDPR 124

RESULT 40
 YPOL ACICA STANDARD; PRT; 204 AA.
 ID YPOL ACICA
 AC P07778;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DS Hypothetical protein in POC-V 5'region (ORF L) (Fragment).
 OS Acinetobacter calcoaceticus.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Moraxellaceae; Acinetobacter.
 OX NCBI_TaxID=471;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=LMD 79.41;
 RC MEDLINE=89123056; PubMed=2536663;
 RX Goosen N., Horsman H.P.A., Huinen R.G.M., van de Putte P.;
 RA "Acinetobacter calcoaceticus genes involved in biosynthesis of the
 RT coenzyme pyrrolo-quinoline-quinone: nucleotide sequence and
 RT expression in Escherichia coli K-12.";
 RL J. Bacteriol. 171:447-455(1989).
 CC -1- FUNCTION: NOT KNOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; X06452; CAA29752.1; -
 DR PIR; A32252; A32252.
 DR InterPro; IPR004843; M-peptidase.
 DR Pfam; PF00149; Metallophos; 1.
 KW Hypothetical protein.
 FT NON TER 204 204
 SQ SEQUENCE 204 AA; 22132 MW; AB5E0D4E7988AB4 CRC64;

Query Match 56.4%; Score 31; DB 1; Length 204;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NKGTQOYT 8
 Db 152 NGGCOOYT 159

Search completed: August 20, 2003, 09:30:42
 Job time : 6.7732 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 20, 2003, 09:19:04 ; Search time 29.0722 Seconds
(without alignments)
88.763 Million cell updates/sec

Title: US-09-991-433-5
Perfect score: 55
Sequence: 1 NKGTQGYTDQ 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23.*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	138	12	Q85155 human parvo
2	55	100.0	141	12	Q85171 human parvo
3	55	100.0	141	12	Q85166 human parvo
4	55	100.0	141	12	Q85146 human parvo
5	55	100.0	142	12	Q85168 human parvo
6	55	100.0	142	12	Q85173 human parvo
7	55	100.0	144	12	Q85138 human parvo
8	55	100.0	145	12	Q85181 human parvo
9	55	100.0	145	12	Q85161 human parvo
10	55	100.0	146	12	Q85158 human parvo
11	55	100.0	147	12	Q85142 human parvo
12	55	100.0	148	12	Q85150 human parvo
13	55	100.0	151	12	Q85131 human parvo
14	55	100.0	151	12	Q85177 human parvo
15	55	100.0	153	12	Q85123 human parvo
16	55	100.0	157	12	Q85196 human parvo

17	55	100.0	162	12	Q85135 human parvo
18	55	100.0	546	12	Q913X0 human parvo
19	55	100.0	546	12	Q913W7 human parvo
20	55	100.0	554	12	Q9PZS9 human parvo
21	55	100.0	554	12	Q90201 human parvo
22	55	100.0	554	12	Q65790 human parvo
23	55	100.0	554	12	Q912B7 human eryth
24	55	100.0	554	12	Q9JGP7 human parvo
25	55	100.0	554	12	Q9WKL9 human parvo
26	55	100.0	554	12	Q8UN54 human parvo
27	55	100.0	760	12	Q9PZT8 human parvo
28	55	100.0	761	12	Q9PZU0 human parvo
29	55	100.0	765	12	Q9PZT6 human parvo
30	55	100.0	769	12	Q9PZT4 human parvo
31	55	100.0	773	12	Q913X1 human parvo
32	55	100.0	773	12	Q913W8 human parvo
33	55	100.0	781	12	Q8UN56 human parvo
34	55	100.0	781	12	Q65789 human parvo
35	55	100.0	781	12	P90223 human parvo
36	55	100.0	781	12	Q85191 human parvo
37	55	100.0	781	12	P90221 human parvo
38	55	100.0	781	12	P89318 human parvo
39	55	100.0	781	12	Q8JYE3 erythroviru
40	55	100.0	781	12	P90224 human parvo
41	55	100.0	781	12	Q9PZT0 human parvo
42	55	100.0	781	12	P89316 human parvo
43	55	100.0	781	12	P89320 human parvo
44	55	100.0	781	12	Q85117 human parvo
45	55	100.0	781	12	P89321 human parvo

ALIGNMENTS

RESULT 1

ID Q85155 PRELIMINARY; PRT; 138 AA.
AC Q85155;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemaier A., Von Pobiotszki A., Giegler A., Cassinotti P., Siegl G.,
RA Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70564; CA94477.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15237 MW; 5899FB8879A3B6BD CRC64;

Query Match 100.0%; Score 55; DB 12; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQGYTDQ 10
DB 107 NKGTQGYTDQ 116

RESULT 2
ID Q85171 PRELIMINARY; PRT; 141 AA.
AC Q85171;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

RESULT 4		
085146		
ID 085146	PRELIMINARY;	PRT; 141 AA.
AC 085146;		
DT 01-NOV-1996	(TREMBLrel. 01,	Created)
DT 01-NOV-1996	(TREMBLrel. 01,	Last sequence update)
DT 01-DEC-2001	(TREMBLrel. 19,	Last annotation update)
DE	Viral protein 1 (fragment).	
GN	VPI.	
OS	Human parvovirus B19.	

RESULT 6		
ID	085173	PRELIMINARY; PRT; 142 AA.
AC	085173;	
DT	01-NOV-1996 (TREMBLrel. 01, Created)	
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	Viral protein 1 (Fragment).	
GN	VP1.	
OS	Human parvovirus B19.	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.	
OX	NCBI_TaxID=10798;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Hemauer A., Von Poblotzki A., Gygler A., Cassinotti P., Siegl G.,...	

RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70582; CA94495.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 142
 SQ SEQUENCE 142 AA; 15820 MW; 3A9E5E07C0BC0434 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 142;
 Best Local Similarity 100.0%; Pred. No. 0.0065;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQ 10
 |||||
 Db 105 NKGTQOYTDQ 114

LT 7
 ID 085138 PRELIMINARY; PRT; 144 AA.

AC 085138;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemaier A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70547; CA94459.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 144
 SQ SEQUENCE 144 AA; 16022 MW; 95D6EF37BA5997AE CRC64;

Query Match 100.0%; Score 55; DB 12; Length 144;
 Best Local Similarity 100.0%; Pred. No. 0.0066;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
 |||||
 Db 107 NKGTQOYTDQ 116

RESULT 8
 Q85181
 ID 085181 PRELIMINARY; PRT; 145 AA.
 AC 085181;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemaier A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70590; CA94503.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.

FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16210 MW; 6B45A1B1B9E923C4 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 0.0067;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQ 10
 |||||
 Db 109 NKGTQOYTDQ 118

RESULT 9
 Q85161
 ID 085161 PRELIMINARY; PRT; 145 AA.

AC 085161;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemaier A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70570; CA94483.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16165 MW; B4FE95AB6E4D9A67 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 145;
 Best Local Similarity 100.0%; Pred. No. 0.0067;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTQOYTDQ 10
 |||||
 Db 105 NKGTQOYTDQ 114

RESULT 10
 Q85158
 ID 085158 PRELIMINARY; PRT; 146 AA.
 AC 085158;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hemaier A., Von Poblitzki A., Gigler A., Cassinotti P., Siegl G.,
 RA Wolf H., Modrow S.;
 RT "XXXXSequence variability among different parvovirus B19 isolates.";
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Z70567; CA94480.1; -;
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 1
 FT NON_TER 146
 SQ SEQUENCE 146 AA; 16252 MW; 2B74FE95AB6E4D9A CRC64;

Query Match 100.0%; Score 55; DB 12; Length 146;

Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 105 NKGTQOYTDQ 114

RESULT 11

QY 085142 PRELIMINARY; PRT; 147 AA.

AC 085142; 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DR 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70551; CAA94463.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 147 147

SEQ SEQUENCE 147 AA; 16424 MW; E4ECF2459B308BA9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 147;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 108 NKGTQOYTDQ 117

RESULT 12

QY 085150 PRELIMINARY; PRT; 148 AA.

AC 085150; 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DR 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70559; CAA94471.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 148 148

SEQ SEQUENCE 148 AA; 16539 MW; 8814ECF2459B308B CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 148;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 109 NKGTQOYTDQ 118

DB 108 NKGTQOYTDQ 117

RESULT 13

QY 085131 PRELIMINARY; PRT; 151 AA.

AC 085131; 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DR 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

SEQUENCE FROM N.A.

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "Sequence variability among different parvovirus B19 isolates.";
RL J. Gen. Virol. 77:1781-1785(1996).

DR EMBL; Z70540; CAA94452.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 151 151

SEQ SEQUENCE 151 AA; 16902 MW; 3CA74914B8E73A3B CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 151;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 109 NKGTQOYTDQ 118

RESULT 14

QY 085177 PRELIMINARY; PRT; 151 AA.

AC 085177; 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DR 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Viral protein 1 (Fragment).

OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

RA Hemauer A., Von Poblotszki A., Giegler A., Cassinotti P., Siegl G.,
Wolf H., Modrow S.;
RT "XXXSequence variability among different parvovirus B19 isolates.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z70586; CAA94499.1; -
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 1
FT NON_TER 151 151

SEQ SEQUENCE 151 AA; 16888 MW; 06C6A614A616DDA3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 151;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
DB 109 NKGTQOYTDQ 118

RESULT 15

Q85123
ID Q85123 PRELIMINARY; PRT; 153 AA.
AC Q85123;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96332516; PubMed=8760426;
RA Hemanuer A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
"Sequence variability among different parvovirus B19 isolates.";
J. Gen. Virol. 77:1781-1785(1996).
EMBL; Z70532; CA94444.1; -;
InterPro; IPR001403; Parvo_coat.
PFam; PF00740; Parvo_coat; 1.
FT NON_TER 1 153
SQ SEQUENCE 153 AA; 17131 MW; ECCALF44020814EC CRC64;
Query Match 100.0%; Score 55; DB 12; Length 153;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
ID Q85123
Db 108 NKGTQOYTDQ 117

RESULT 16
ID Q85196 PRELIMINARY; PRT; 157 AA.
AC Q85196;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
"XXXSequence variability among different parvovirus B19 isolates.";
Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70603; CA94518.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 157
SQ SEQUENCE 157 AA; 17591 MW; 1D42191887FFCE03 CRC64;
Query Match 100.0%; Score 55; DB 12; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
ID Q85135
Db 110 NKGTQOYTDQ 119

RESULT 17
ID Q85135 PRELIMINARY; PRT; 162 AA.
AC Q85135;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Viral protein 1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemanuer A., Von Poblitzki A., Giegler A., Cassinotti P., Siegl G.,
Wolff H., Modrow S.;
"XXXSequence variability among different parvovirus B19 isolates.";
Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z70544; CA94456.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 1 162
SQ SEQUENCE 162 AA; 18005 MW; 3F51443566660F2 CRC64;
Query Match 100.0%; Score 55; DB 12; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
ID Q913X0
Db 109 NKGTQOYTDQ 118

RESULT 18
ID Q913X0 PRELIMINARY; PRT; 546 AA.
AC Q913X0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major virus capsid protein VP2 (Fragment).
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hekynar K., Soderlund-Venemo M., Ranki A., Kivlinoto O., Partio E.K.,
Hedman K.;
"A new parvovirus B19 genotype persistent in skin.";
Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AK95572.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 546 546
SQ SEQUENCE 546 AA; 59934 MW; 467BE468A67282B5 CRC64;
Query Match 100.0%; Score 55; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
ID Q913W7
Db 395 NKGTQOYTDQ 404

RESULT 19
ID Q913W7 PRELIMINARY; PRT; 546 AA.
AC Q913W7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major virus capsid protein VP2 (Fragment).
GN VP2.
OS Human parvovirus B19.

```
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAM;
RA Hokyar K., Soderlund-Venemo M., Ranki A., Kiviluoto O., Partio E.K.,
  Hedman K.;
RT "A new parvovirus B19 genotype persistent in skin.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AI044268; AAK95575.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
DR NON_TER 546 546
FT SEQUENCE 546 AA; 59910 MW; BBF1DD23695E3FCD CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 546;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
|||||
395 NKGTQOYTDQ 404

RESULT 20
Q9PZS9 PRELIMINARY; PRT; 554 AA.
AC Q9PZS9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP2 capsid protein.
GN VP.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HV;
RA Gallinella G., Venturoli S.;
RT "B19 Genome Sequence and Structure Analysis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF162273; AAD4615.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60833 MW; 7901FEBD65A697E6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
|||||
395 NKGTQOYTDQ 404

RESULT 21
Q90201 PRELIMINARY; PRT; 554 AA.
AC Q90201;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP2 structural protein (Capsid protein VP2) (Fragment).
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9708118; PubMed=8922470;
RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
```

```
RT VP1/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MI, and N6;
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K.,
  Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with
  rheumatoid arthritis.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=N8;
RX MEDLINE=90218047; PubMed=2157807;
RA Umene K., Nunoe T.;
RT "The genome type of human parvovirus B19 strains isolated in Japan
  during 1981 differs from types detected in 1986 to 1987: a correlation
  between genome type and prevalence.";
RL J. Gen. Virol. 71:983-986 (1990).
DR EMBL: U53595; AAB47453.1; -.
DR EMBL: U53596; AAB47455.1; -.
DR EMBL: U53593; AAB47449.1; -.
DR EMBL: U53594; AAB47451.1; -.
DR EMBL: U53597; AAB47457.1; -.
DR EMBL: U53600; AAB47463.1; -.
DR EMBL: U53601; AAB47465.1; -.
DR EMBL: U53598; AAB47459.1; -.
DR EMBL: U53589; AAB47461.1; -.
DR EMBL: AB030693; BAA90290.1; -.
DR EMBL: AB030673; BAA90268.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
DR NON_TER 554 554
FT SEQUENCE 554 AA; 60853 MW; 98FE598FF20CB66F CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
|||||
395 NKGTQOYTDQ 404

RESULT 22
Q65790 PRELIMINARY; PRT; 554 AA.
AC Q65790;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VEN1;
RA Echavarría Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U1358; AAB3559.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR Pfam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60798 MW; BA89F2B293BB4E24 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
|||||
395 NKGTQOYTDQ 404
```

```

RESULT 23
ID Q912B7 PRELIMINARY; PRT; 554 AA.
AC Q912B7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
GN VP2 protein.
OS Human erythrovirus V9.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=72197;
RN [1]
RP STRAIN=V9.
RC SEQUENCE FROM N.A.
RA Nguyen O.T.;
RT "Molecular cloning and sequencing of a novel human erythrovirus genome: new species beside B19 in the genus Erythrovirus."
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AJ249437; CAC80622.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60885 MW; B4F338CDA80F336 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404

RESULT 24
ID Q9UGP7 PRELIMINARY; PRT; 554 AA.
AC Q9UGP7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP STRAIN=RM;
RC SEQUENCE FROM N.A.
RA Ishii K.K., Munakata Y., Funato T., Fu Y., Koseki N., Sugamura K., Sasaki T.;
RT "Sequence of human parvovirus B19 isolates from patients with rheumatoid arthritis."
Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AB030694; BAA90293.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60839 MW; 0DB958B33C73EF64 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404

RESULT 25
ID Q9WKL9 PRELIMINARY; PRT; 554 AA.
AC Q9WKL9;
DT 01-NOV-1999 (TRENBLrel. 12, Created)

```

```

DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Structural protein VP2.
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20014169; PubMed=10548133.
RA Hemaier A., Beckenlehner K., Wolf H., Lang B., Modrow S.;
RT "Acute parvovirus B19 infection in connection with a flare of systemic lupus erythematosus in a female patient."
J. Clin. Virol. 14:73-77(1999).
DR EMBL; AF113323; AAC99439.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60913 MW; 98FB588FE205C66F CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404

RESULT 26
ID Q8JNS4 PRELIMINARY; PRT; 554 AA.
AC Q8JNS4;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
GN VP2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;
RN [1]
RP STRAIN=D91.1;
RC MEDLINE=22174902; PubMed=12186896;
RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G., Merlier J.F., Garbarg-Chenon A.;
RT "Genetic Diversity within Human Erythroviruses: Identification of Three Genotypes."
J. Virol. 76:9124-9134(2002).
DR EMBL; AY083234; AAL91014.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 554 AA; 60830 MW; A298575C083F24C8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 554;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTQOYTDQ 10
Db 395 NKGTQOYTDQ 404

RESULT 27
ID Q9PZT8 PRELIMINARY; PRT; 760 AA.
AC Q9PZT8;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
GN VP1/2 (Fragment).
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
NCBI_TaxID=10798;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kat12;
 RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venermo M., Kiviluoto O.,
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025(2000).
 DR EMBL; AF161224; AAD45912.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 760
 SQ SEQUENCE 760 AA; 83403 MW; EA6EE0145E3A0E5A CRC64;

Query Match 100.0%; Score 55; DB 12; Length 760;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTOQYTDQ 10
 |||||
 622 NKGTOQYTDQ 631

RESULT 28

G9PZU0 PRELIMINARY; PRT; 761 AA.

ID G9PZU0
 AC G9PZU0:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE VPI/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Kat11;
 RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venermo M., Kiviluoto O.,
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025(2000).
 DR EMBL; AF161223; AAD45910.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 761
 SQ SEQUENCE 761 AA; 83574 MW; 73A84C930D473530 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 761;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
 |||||
 DB 622 NKGTOQYTDQ 631

RESULT 29

G9PZT6 PRELIMINARY; PRT; 765 AA.

ID G9PZT6
 AC G9PZT6:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-DEC-2001 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE VPI/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kat13;

RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venermo M., Kiviluoto O.,
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025(2000).
 DR EMBL; AF161225; AAD45915.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 765
 SQ SEQUENCE 765 AA; 83998 MW; 89E2546086DCDBE8 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 765;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
 |||||
 DB 622 NKGTOQYTDQ 631

RESULT 30

G9PZT4 PRELIMINARY; PRT; 769 AA.

ID G9PZT4
 AC G9PZT4:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE VPI/2 (Fragment).
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kat14;
 RX MEDLINE=20191963; PubMed=10725428;
 RA Hokynar K., Brunstein J., Soderlund-Venermo M., Kiviluoto O.,
 RT "Integrity and full coding sequence of B19 virus DNA persisting in
 human synovial tissue."
 RL J. Gen. Virol. 81:1017-1025(2000).
 DR EMBL; AF161226; AAD45917.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 FT NON_TER 769
 SQ SEQUENCE 769 AA; 84578 MW; 0749D46E5CA7BB68 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 769;
 Best Local Similarity 100.0%; Pred. No. 0.04;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKGTOQYTDQ 10
 |||||
 DB 622 NKGTOQYTDQ 631

RESULT 31

G913X1 PRELIMINARY; PRT; 773 AA.

ID G913X1
 AC G913X1:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Minor virus capsid protein VPI (Fragment).
 GN VPI.

OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.

OX NCBI_TaxID=10798;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Lali;
 RA Hokynar K., Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K.,
 Hedman K.;

RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95574.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85164 MW; D60CCA6F90B05378 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
|||
Db 622 NKGTQOYTDQ 631

RESULT 32

OQ913W8 PRELIMINARY; PRT; 773 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Minor virus capsid protein VP1 (Fragment).
GN VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAM;
RA Hokynar K., Soderlund-Venermo M., Ranki A., Kiviluoto O., Partio E.K.,
RT "A new parvovirus B19 genotype persistent in skin."
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY044266; AAK95574.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
FT NON_TER 773
SQ SEQUENCE 773 AA; 85140 MW; 2886F3245F9CE850 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NKGTQOYTDQ 10
|||
622 NKGTQOYTDQ 631

RESULT 33

OQ8UN56 PRELIMINARY; PRT; 781 AA.

AC OQ8UN56;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D91.1;
RX MEDLINE=22174902; PubMed=12186896;
RA Servant A., Laperche S., Lallemand F., Marinho V., De Saint Maur G.,
RT Merlet J.F., Garbarg-Chenon A.;
RT "Genetic Diversity within Human Erythroviruses: Identification of
RT Three Genotypes."
RL J. Virol. 76:9124-9134(2002).
DR EMBL; AY083234; AAL91013.1; -.
DR InterPro; IPR001403; Parvo_coat.

DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86142 MW; 4EB71FEDD41FC8F3 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
|||
Db 622 NKGTQOYTDQ 631

RESULT 34

O65789 PRELIMINARY; PRT; 781 AA.

AC O65789;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VP1.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VEN1;
RA Echevarria Mayo J.E., Erdman D.D.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; U31358; AAA83558.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 85985 MW; E1D5FDE230935DDO CRC64;

Query Match 100.0%; Score 55; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
|||
Db 622 NKGTQOYTDQ 631

RESULT 35

P90223 PRELIMINARY; PRT; 781 AA.

AC P90223;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USA2;
RX MEDLINE=97081188; PubMed=8922470;
RA Erdman D.D., Durigton E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates."
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; U38514; AAB47796.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; C1D911E39CF04A88 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NKGTQOYTDQ 10
|||
Db 622 NKGTQOYTDQ 631

```

RESULT 36
ID Q85191 PRELIMINARY; PRT; 781 AA.
AC Q85191;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF2.
GN ORF2.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA Hemauer A., Von Podlitzki A., Gigler A., Casinotti P., Siegl G.,
RA Wolf H., Modrow S.;
RT "XXXsequence variability among different parvovirus B19 isolates.";
EMBL; Z70599; CA94513.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86026 MW; AA02577B683EDB2D CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 781;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 NKGTQOYTDQ 10
DB 622 NKGTQOYTDQ 631

RESULT 37
ID P90221 PRELIMINARY; PRT; 781 AA.
AC P90221;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=KOR2;
RA MEDLINE=97081188; PubMed=8922470;
RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates.";
EMBL; U38511; AAB4793.1; -.
InterPro: IPR001403; Parvo_coat.
PFam: PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86070 MW; 9FA830083F6F1357 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 781;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 NKGTQOYTDQ 10
DB 622 NKGTQOYTDQ 631

RESULT 38
ID P89318 PRELIMINARY; PRT; 781 AA.
AC P89318;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

```

```

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=JAP1;
RA MEDLINE=97081188; PubMed=8922470;
RA Erdman D.D., Durigson E.L., Wang Q.Y., Anderson L.J.;
RT "Genetic diversity of human parvovirus B19: sequence analysis of the
RT VP1/VP2 gene from multiple isolates.";
RL J. Gen. Virol. 77:2767-2774(1996).
DR EMBL; U38509; AAB4791.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR PFam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86064 MW; 3AE65CE89096339 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 781;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 NKGTQOYTDQ 10
DB 622 NKGTQOYTDQ 631

RESULT 39
ID Q8UYE3 PRELIMINARY; PRT; 781 AA.
AC Q8UYE3;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE 7.5 KDa protein.
OS Erythrovirus A6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=182494;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen Q.T., Wong S., Brown K.E.;
RT "Identification and characterization of a second novel human
RT erythrovirus variant, A6.";
RT Virology 0:0-0(2002).
DR EMBL; AY064475; AAL55418.1; -.
DR InterPro: IPR001403; Parvo_coat.
DR PFam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 781 AA; 86277 MW; E73AB663B4551128 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 12; Length 781;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 NKGTQOYTDQ 10
DB 622 NKGTQOYTDQ 631

RESULT 40
ID P90224 PRELIMINARY; PRT; 781 AA.
AC P90224;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VP1 and VP2 structural protein.
OS Human parvovirus B19.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
OX NCBI_TaxID=10798;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BR21;
RA MEDLINE=97081188; PubMed=8922470;

```

RA Erdman D.D., Durigon E.L., Wang Q.Y., Anderson L.J.;
 RT "Genetic diversity of human parvovirus B19: sequence analysis of the
 VPI/VP2 gene from multiple isolates."
 RL J. Gen. Virol. 77:2767-2774 (1996).
 DR EMBL; U38546; AAB47801.1; -.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 SQ SEQUENCE 781 AA; 86039 MW; C3A29EB1DB8AD378 CRC64;

Query Match 100.0%; Score 55; DB 12; Length 781;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NKGTQGYTDQ 10
 |||||
 Db 622 NKGTQGYTDQ 631

Search completed: August 20, 2003, 09:23:59
 Elapsed time: 30.0722 secs

THIS PAGE BLANK (USPTO)